TEC-0123

Research in Knowledge-Based Automatic Feature Extraction

R. Nevatia
A. Huertas

University of Southern California Institute for Robotics and Intelligent Systems
Powell Hall, 204
Los Angeles, CA 90089-0273

May 1999

Prepared for:
Defense Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington, VA 22203-1714

Monitored by:
U.S. Army Corps of Engineers
Topographic Engineering Center
7701 Telegraph Road
Alexandria, Virginia 22315-3864

US Army Corps of Engineers
Topographic Engineering Center

T

E

C

Destroy this report when no longer needed.
Do not return it to the originator.

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

The citation in this report of trade names of commercially available products does not constitute official endorsement or approval of the use of such products.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 1999	Technical 23 March 1997 - 31 March 1998

4. TITLE AND SUBTITLE

Research in Knowledge-Based Automatic Feature Extraction

5. FUNDING NUMBERS

DACA76-97-K-0001

6. AUTHOR(S)

R. Nevatia and A. Huertas

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Southern California
Institute for Robotics and Intelligent Systems
Powell Hall, 204
Los Angeles, CA 90089-0273

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Defense Advanced Research Projects Agency
3701 N. Fairfax Drive, Arlington, VA 22203-1714

U.S. Army Topographic Engineering Center
7701 Telegraph Road, Alexandria, VA 22315-3864

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

TEC-0123

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report provides a description of USC's activities as part of the DARPA Automated Population of Geospatial Databases (APGD) project. Focus of this effort is on detection, delineation, and description of 3-D buildings from aerial images. Buildings are objects of obvious importance in urban environments and accurate models of them are needed for a number of tasks such as mission planning, mission rehearsal, tactical training, damage assessment, and change detection. Difficulties in modeling buildings come from problems of segmentation, 3-D inference, and shape description. Development of a multi-image-based system is described. This system operates on a hypothesize and verify paradigm. Hypotheses for rectilinear rooftops are generated by hierarchical grouping of lower level features. Three-dimensional cues from walls and shadows are used to verify and select from these hypotheses. Three-dimensional cues from range sensors, such as IFSAR, can be used when available. Results on selected windows of Fort Hood, TX, and Fort Benning, GA, sites are presented. An efficient interactive system can be used to edit the results of the automatic system to reduce the effort required from the user. The current system is limited to rectilinear-shaped buildings and assumes that ground height, camera models, and illumination parameters are known.

14. SUBJECT TERMS

3-D model construction, building detection, sensor fusion

15. NUMBER OF PAGES

55

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

TABLE OF CONTENTS

ILLUSTRATIONS

Acknowledgments

The work reported here is the result of contributions from several faculty, research staff and graduate students, under the direction of Prof. Ram Nevatia. The principal developers were Sanjay Noronha (baseline multiview system), ZuWhan Kim (multiview system upgrades, analysis tools and Bayesian network classifiers), Andres Huertas (multiview system upgrades, shadow and wall evidence, IFSAR integration, gable roof analysis, batch system and its interface to the SOCET Set system of GDE Systems, San Diego, incorporation of data sets into RCDE) and Jian Li (interactive system). Prof. Keith Price assisted in all aspects related to the use of the RCDE and LISP programming in general, in testing of the system, and in porting updated batch versions of the systems to GDE Systems.

Preface

This research was sponsored by the Defence Advanced Research Projects Agency (DARPA) and monitored by the U.S. Army Topographic Engineering Center (TEC) under contract DACA76-97-K-0001, titled "Research in Know-ledge-Based Automatic Feature Extraction." The DARPA Program Managers is Mr. George Lukes, and the TEC Contracting Officer's Representative is Ms. Lauretta Williams.

1 Introduction and Overview

This is a technical report describing the activities of the University of Southern California under contract # DACA76-97-K-0001, during the period of March 23, 1997 to March 31, 1998. The goal of this project is to develop methods for automatic feature extraction for population of geospatial databases (APGD) with particular focus on buildings. The task includes detection, delineation and description of 3-D buildings. Buildings are objects of obvious importance in urban environments and accurate models for them are needed for a number of battlefield awareness tasks such as mission planning, mission rehearsal, tactical training and damage assessment. Other applications include intelligence analysis for site monitoring and change detection.

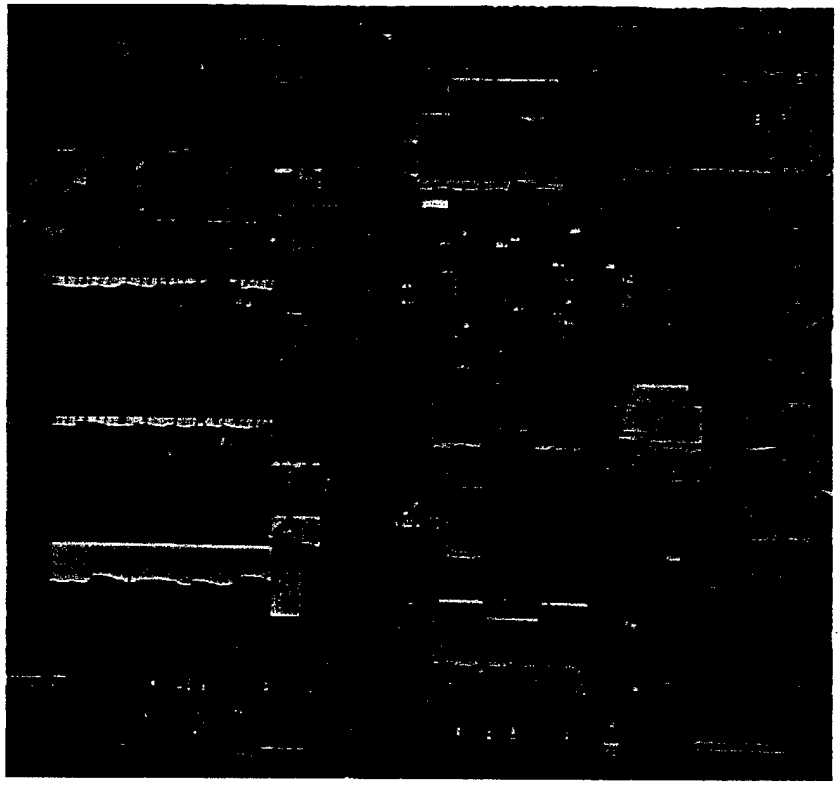
The problem of 3-D feature extraction is difficult in many ways. Low level segmentation techniques (such as line detection) give incomplete and imperfect results. Object boundaries may be incomplete and many extraneous boundaries are present. As example, Figure 1.1 (a) shows a window from an image of the Fort Hood, TX, site. Humans can readily see the buildings (and other structures in it) but the lines detected from this image shown in Figure 1.1 (b) show the difficulties that an automated process needs to overcome. The building boundaries are fragmented and boundaries from many other sources such as roads, sidewalks, landscaping, trees and cars are present. The system must complete the desired object boundaries and discard the extraneous ones.

The second difficulty is in inferring the 3-D structure of the object as this information is not explicit in an intensity image. In fact, the segmentation and 3-D inference problems are highly related; 3-D information makes the segmentation task easier by distinguishing between surface and illumination boundaries, and it is easier to extract 3-D shape from segmented objects. In recent years interferometric synthetic aperture radar (IFSAR) sensors that directly infer 3-D positions of visible points have started to become available. Use of such sensors allows discrimination of surface discontinuities. However, some extraneous boundaries due to noise and other sources remain. Also, IFSAR data typically has areas of missing elements and may contain some points with grossly erroneous values. Chapter 2 describes use of IFSAR in the task of automatic building detection.

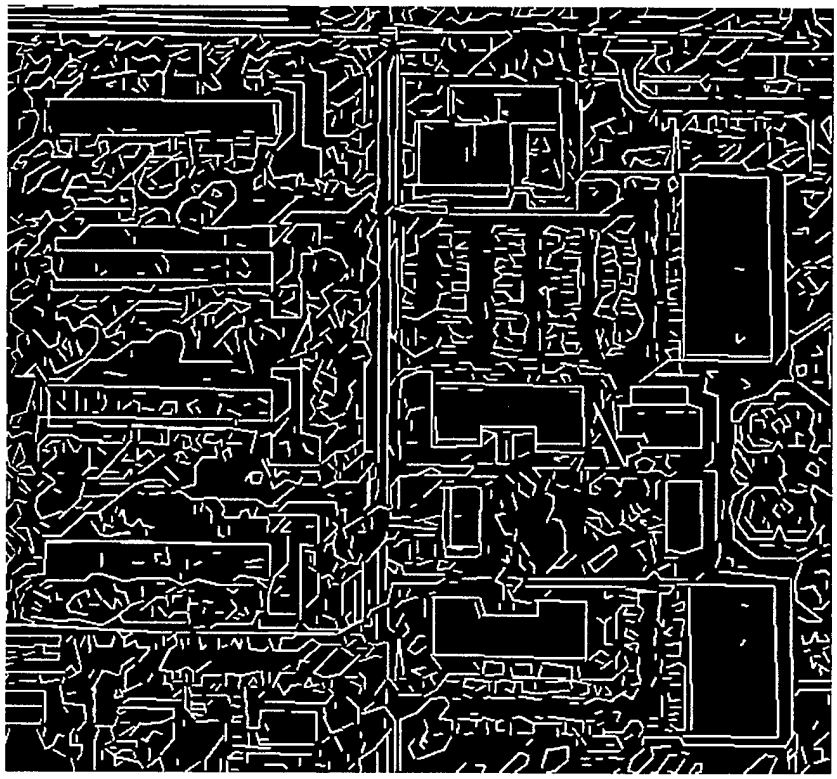
In the absence of an ideal range sensor, depth can be inferred from multiple images. However, this requires finding corresponding points or features in two or more images, which is a complex task in itself. Figure 1.2 (a) shows a second view of the scene shown earlier in Figure 1.1 (a), and Figure 1.2 (b) shows the lines detected in it. Clearly, the task of corresponding points and lines in the two views is difficult, at least by using local information alone and some grouping of such features is required for 3-D inference.

Once objects have been segmented and 3-D shape recovered, the task of shape description still remains. This consists of forming complex shapes from simpler shapes that may be detected at earlier stages. For example, a building may have several wings, possibly of different heights, that may be detected as separate parts rather than one structure initially.

The approach used in this effort is to use a combination of tools: reconstruction and reasoning in 3-D, use of multiple sources of data and perceptual grouping as shown in Figure 1.3. Context and domain knowl-

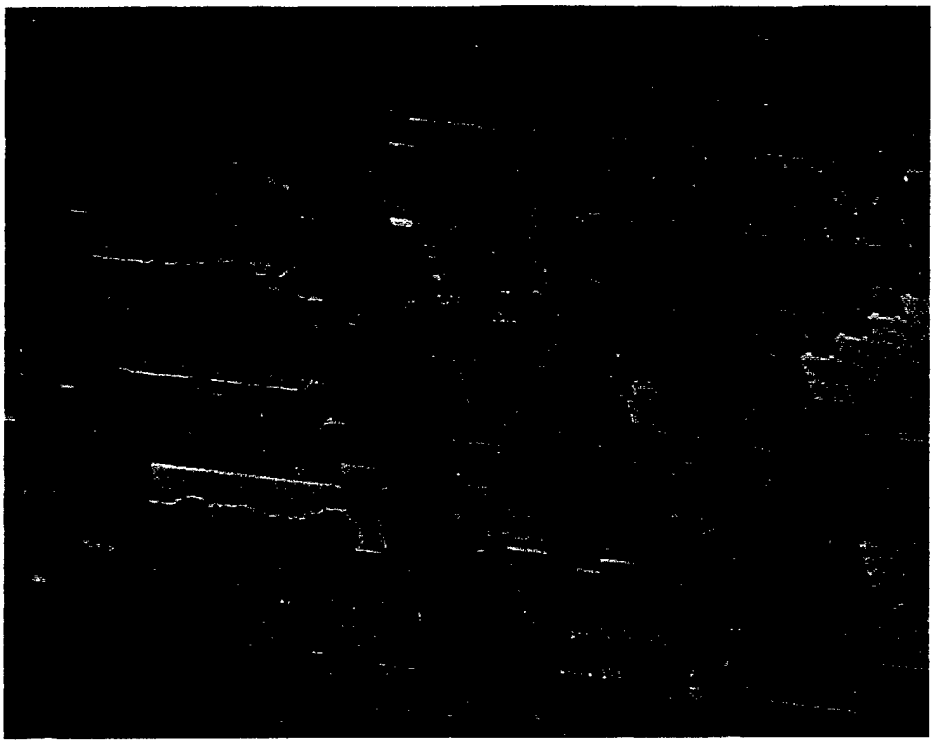


(a)

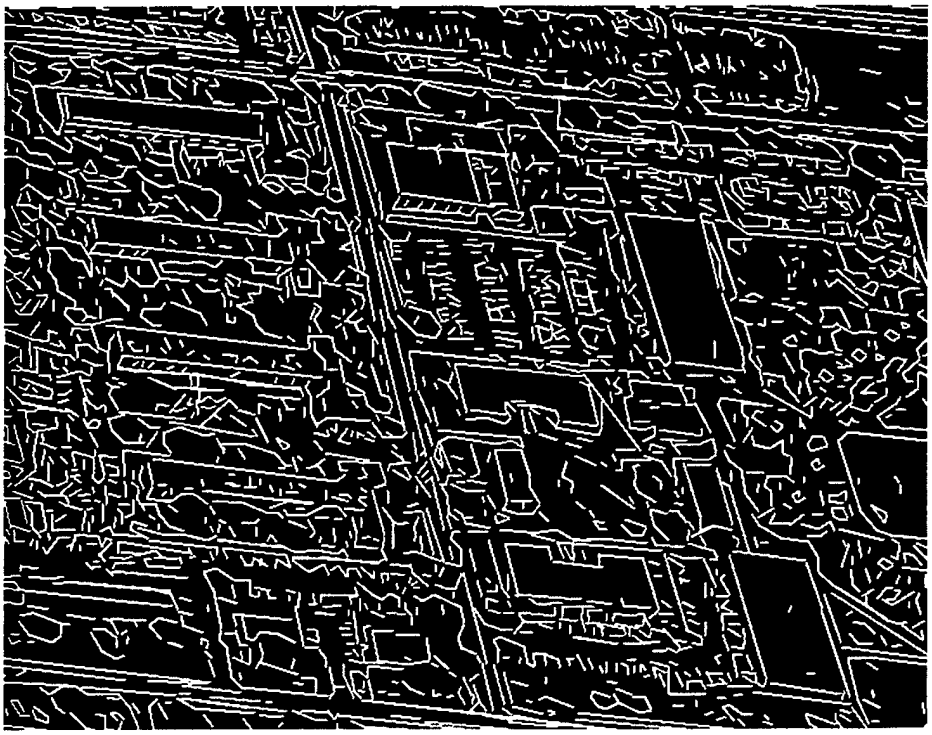


(b)

Figure 1.1 (a) A window from the 'headquarters' area of Fort Hood; (b) Detected lines



(a)



(b)

Figure 1.2 (a) Another window from the 'headquarters' area; (b) Detected lines

edge guide the applications of these tools. Context comes from knowledge of camera parameters, geometry of objects to be detected and illumination conditions (primarily the sun position). Some knowledge of the approximate terrain is also utilized. The information form sensors of different modalities, such as IFSAR and EO (electro-optical) images is fused not at pixel level but at higher feature levels.

The described system is limited to buildings with rectilinear roofs; this results in rooftops projection to parallelograms or combinations thereof. Most of the work has been on buildings with flat roofs but work on slanted roofs has been initiated. It is assumed that camera models are given and approximated by orthographic projection locally, that the ground is flat with known height and that the sun position is given (computable from latitude, longitude and time of day data). Multiple images are *not* assumed to have been taken at the same time. The details of this system are given in Chapter 3.

Even though the long-term goal of this effort is complete automation, it is likely that the systems that can be developed in the near term will not be perfect and will miss some objects or find incorrect ones. However, it is possible to use partial results of the automatic analysis or to improve the automatic analysis by providing some user input. The aim is to make this user input efficient, requiring much less effort than would be necessary for conventional interactive systems which largely take care only of geometric computations and bookkeeping. Efforts towards developing such interactive strategies is described in Chapter 4.

Chapter 5 provides a number of results of this system and topics for future research.

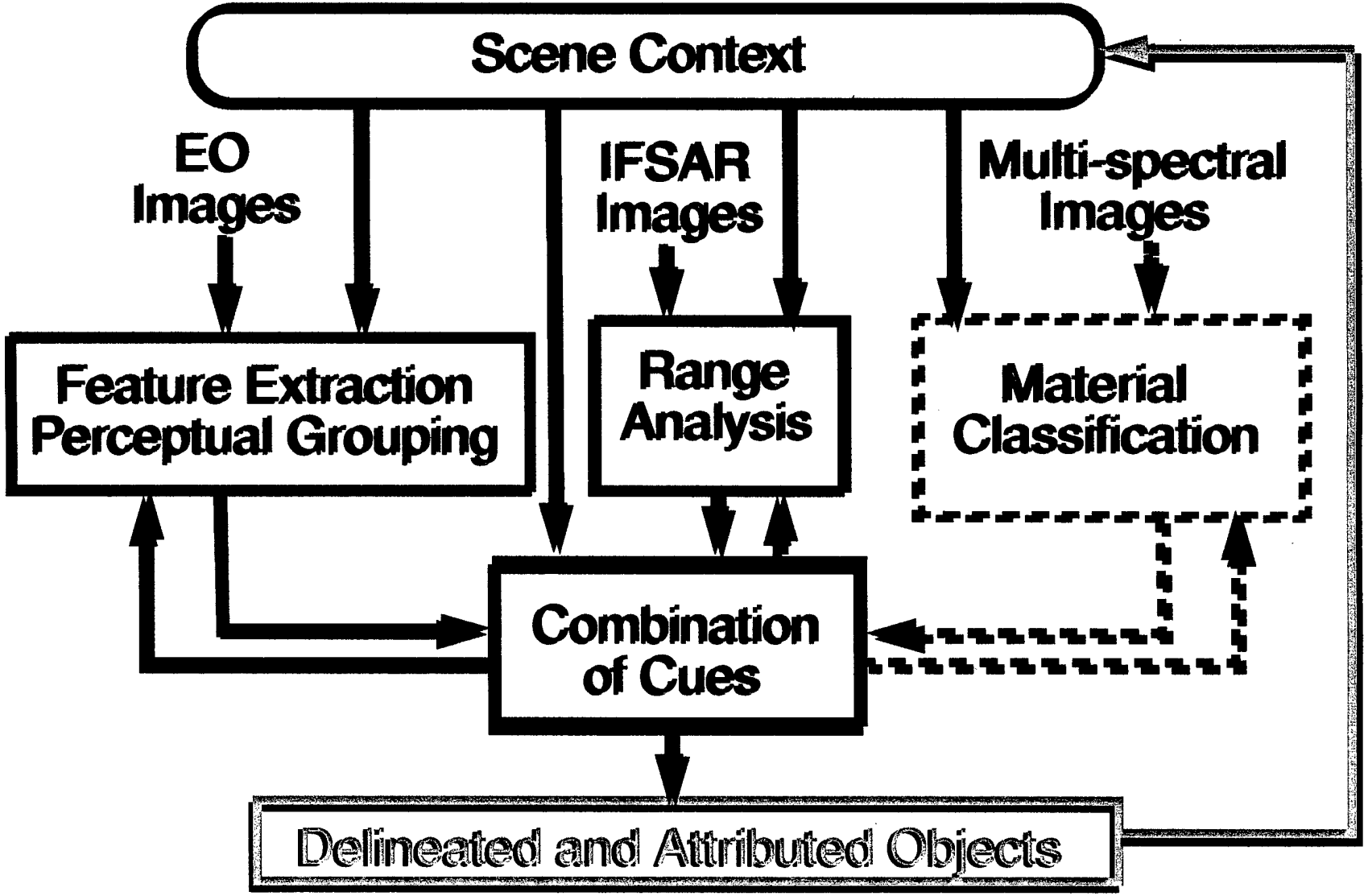


Figure 1.3 Block diagram of the system

2 Cues From IFSAR

IFSAR data, in addition to reflectivity information, also contains information of 3-D points in a scene. The 3-D information allows cues for presence of buildings to be extracted more easily than for panchromatic images. However, the resolution of the IFSAR images is more limited and many wrong values are present. Thus, it is preferable to use IFSAR for detection and panchromatic images for accurate delineation. Cues from IFSAR data can be used in a number of ways: in selecting areas to process where buildings may be present, in eliminating certain building hypotheses, and in adding confidence to the presence of buildings.

IFSAR data is given in the form of three images, called the **mag**, **dte** and **cor** images. The **mag** image is like a normal intensity image measuring the amount of reflected signal coming back to the sensor. The **dte** image encodes the 3-D information in the form of a digital terrain elevation map where the pixel values define the height of the corresponding scene point. The **cor** image contains the phase correlation information between two images used for the interferometric process; it can be useful in distinguishing among types of materials as the returns associated with objects that remain stationary are highly correlated.

Two data sets have been available for testing. One is over the Fort Hood, TX site and has ground resolution of 2.5 meter (uses the IFSARE sensor). The other is over the McKennena MOUT site at Fort Benning, GA with a ground resolution of 0.4 meters obtained by the Scandia searchlight IFSAR process. The two data collection modes have different characteristics and require slightly different modes of processing.

Although the primary source of cues for buildings appears to be the **dte** image, an initial characterization of the data indicates that at low resolutions it is appropriate to extract cues that indicate the possible presence of significant features, such as buildings and trees, from a combination of the **mag**, **dte** and **cor** images. Consider the portion of an image from the Fort Hood site shown in Figure 2.1. It contains 12 buildings. The corresponding 2.5 meter **mag**, **dte** and **cor** images are shown in Figure 2.2. Only the three top buildings appear salient in the **dte** image but all the buildings are well represented in the **mag** image.

Clearly the orientation of the buildings with respect to the direction of flight during acquisition contributes to the resulting returns, and some buildings may not be adequately represented for some orientations. Thus, it is advantageous to use a combination of the **dte** and **mag** images with a higher weight given to the **dte** contribution. The phase correlation image is used to take advantage of the behavior of the signal phase in the IFSAR process and thus help support cues for buildings derived from **mag** and/or **dte** components taken individually or in combination.

Figure 2.3 shows that, at low resolutions, it is not sufficient to threshold the **dte** images to obtain cues corresponding to objects of interest. Figure 2.3 (a) shows the **dte** regions "just above the ground", that is, about 1.5 meters above the ground (mean elevation). Figure 2.3 (b) shows the thresholded **dte** image at the mean intensity plus one standard deviation (5.0 meters). The buildings are somewhat apparent in this image but the presence of many artifacts would be misleading to an automated system beyond a rough indication of possible presence of a building.

Instead, the **mag** and **dte** images are processed by applying a technique similar to the one described in [Chen, et al., 1987]. Regions of interest are extracted from the **mag** and **dte** images correspond to the positive-valued regions in the output of the convolution of the images with a Laplacian-of-Gaussian (LOG) filter. These regions are shown in Figure 2.4 (a) and Figure 2.4 (b) for the **mag** and **dte** images respectively. These images are combined linearly with a weight of 10 given to the **dte** regions. The resulting image is then thresholded to determine the regions of interest. The *cues* image, shown in Figure 2.4 (c) is computed by selecting those regions that have a high correlation component. Figure 2.4(d) shows the connected components that have a certain minimum size (area) and taken to correspond to cues for building structures and other tall objects.

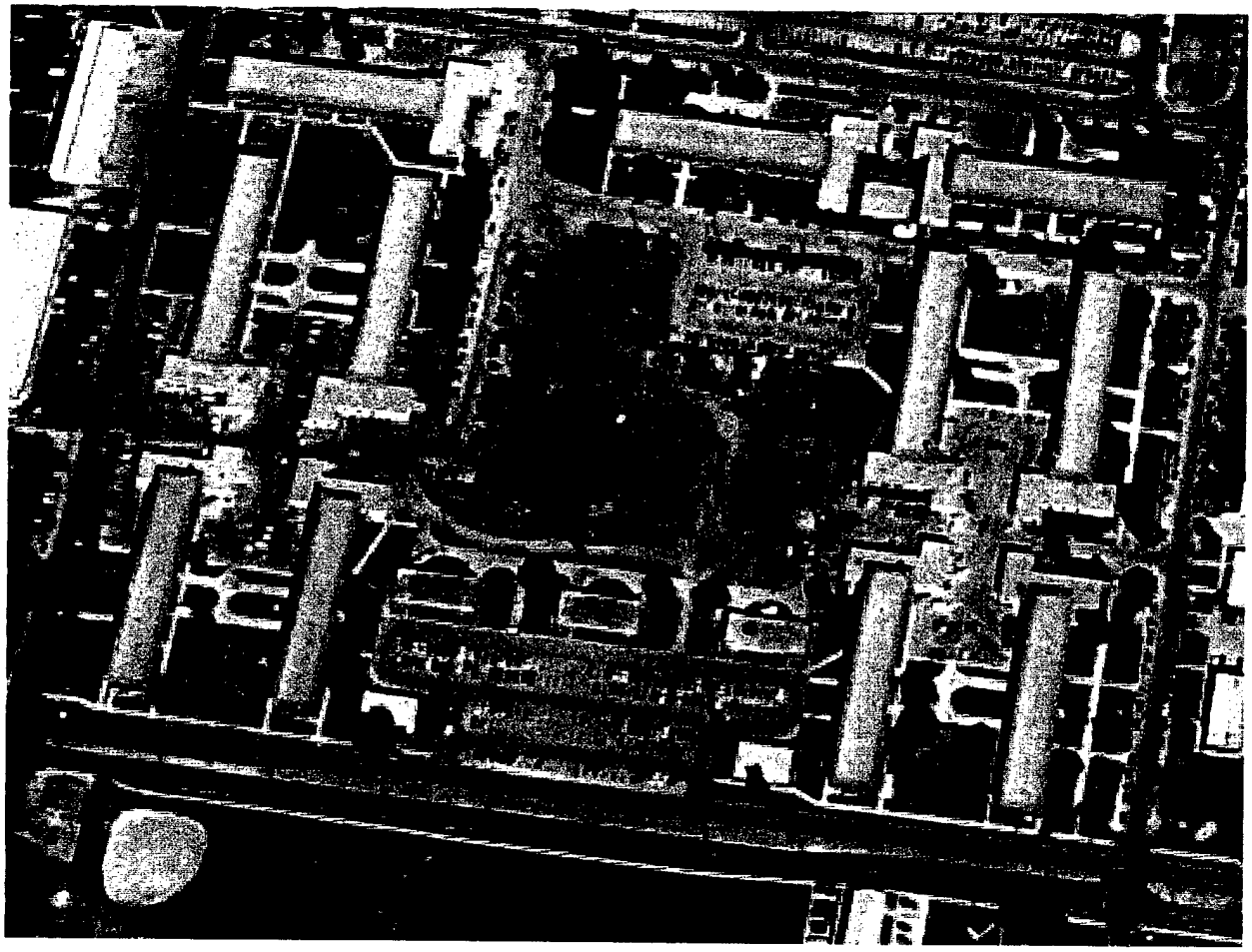


Figure 2.1 Portion of an image from the Fort Hood site (Admin. Area 3)

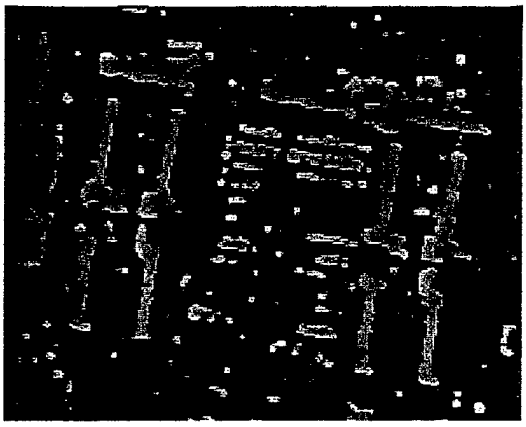
The detected cue regions are shown approximated by bounding rectangles in Figure 2.5 (a). Figure 2.5 (b) show the cues projected on the panchromatic image shown earlier in Figure 2.1. The sensor model for the IFSAR view was computed by approximating an overhead viewing angle. The "height" of the cue rectangles for this projection was determined form the underlying **dte** data.

The camera model or sensor transforms needed to correspond features in the electro-optical (EO) images and in the radar images was derived manually by approximating a central projection camera with an overhead viewpoint. The position of the sensor was derived by incorporating the radar images as 2-D worlds into the 3-D reference coordinate system of the EO imagery as defined by the RCDE system. The

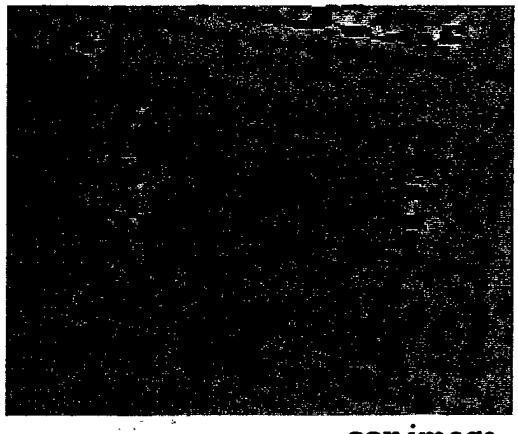
principal points in the radar images was made to correspond to the center of the images, and the sensor "altitude" was manually adjusted to obtain a reasonable scaling transform.

Figure 2.6 through Figure 2.10 show another area from Fort Hood, and the corresponding steps and results in the extraction of useful cues from the available radar data. The same parameters were used for this example as were used for the example given above.

The radar dataset for the McKenna MOUT site is derived by a different process using four separate passes and has different characteristics than the data from the Fort Hood set. The **mag** and **cor** images, shown in Figure 2.11, do not seem to have easily usable characteristics for reliable building cueing. However, the **dte** image, shown on the left in Figure 2.12 exhibits good characteristics. The IFSAR cues, shown on the right, are derived from the **dte** image alone to represent the rough foot-prints of the elevated objects, such as buildings and trees. The regions shown in Figure 2.12 are derived as described earlier, by convolution with a Laplacian-of-Gaussian filter that smooths the image and locates the object boundaries by the positive-valued regions bounded by the zero-crossings in the convolution output. Figure 2.13 shows that all the buildings, their approximate location, size and orientation, are well represented. The baseline building models provided by SRI International of Menlo Park, CA are shown projected on the cue regions derived from the **dte** image. The nearby trees on the South and West sides of the site are also well represented.



mag image



cor image

dte image

Figure 2.2 Magnitude, Elevation and Correlation images for Admin Area 3

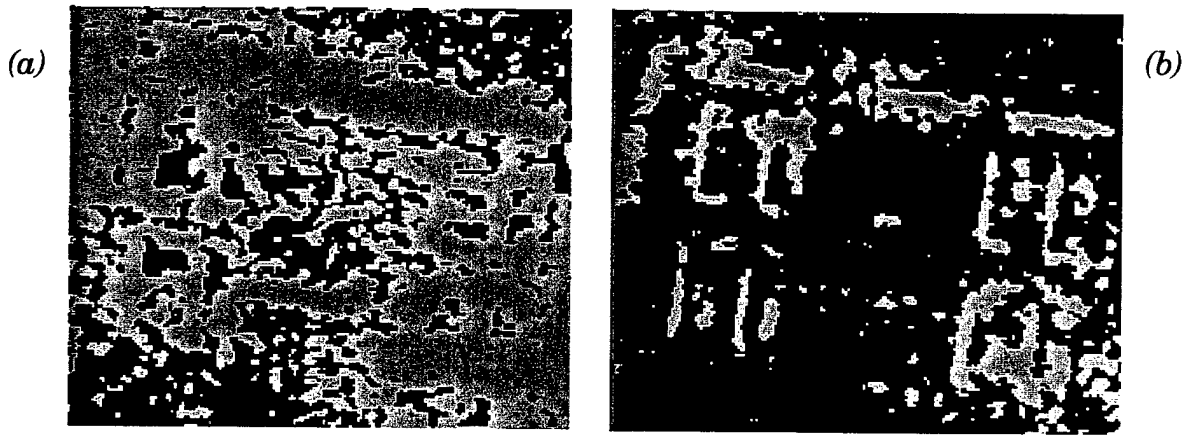


Figure 2.3 Simple thresholding of dte image is not appropriate to derive cues: (a) thresholded at mean elevation; (b) thresholded at mean plus one standard deviation

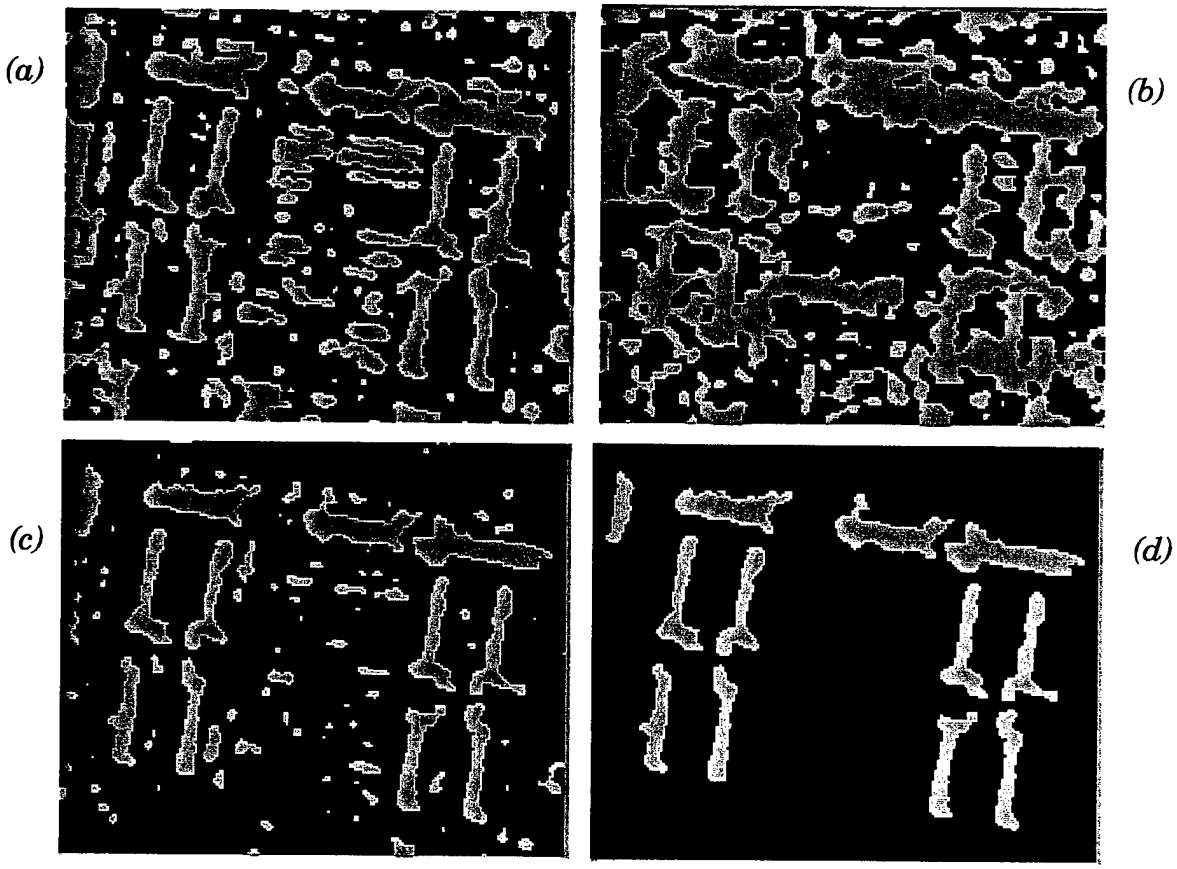


Figure 2.4 (a) Positive LOG convolution regions from mag image; (b) Positive LOG convolution regions from dte image; (c) Combination of (a) and (b) with high phase correlation (> 0.9); (d) Connected components of (c) denote building cues

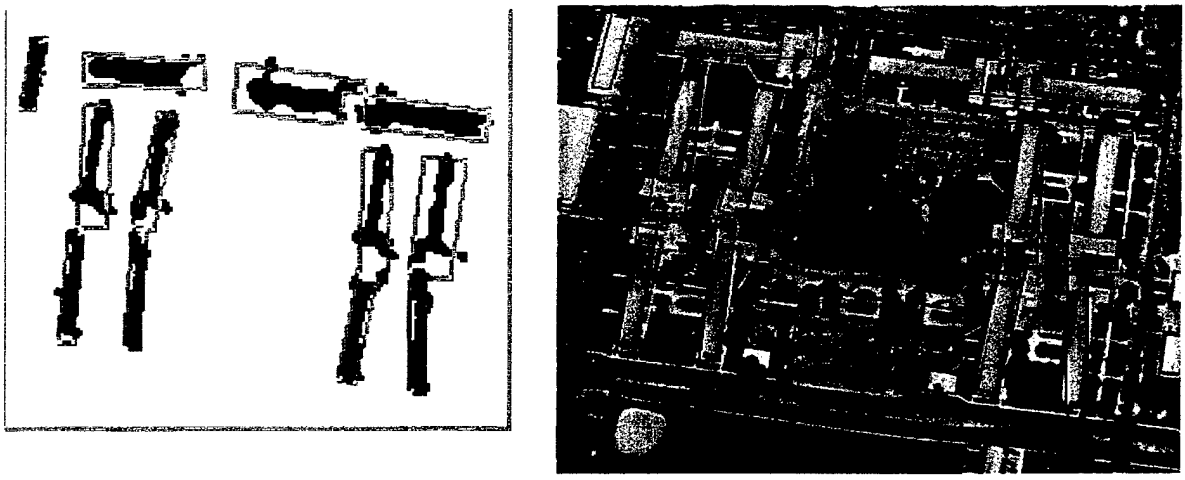


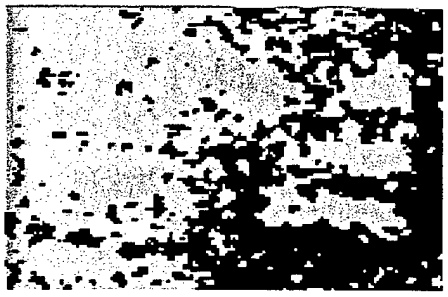
Figure 2.5 (a) IFSAR cues approximated by bounding rectangles and (b) projected on panchromatic image at 10 meters elevation

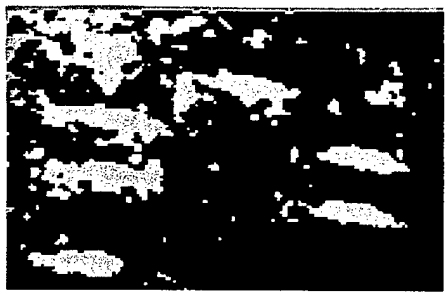
Figure 2.6 Another portion of an image from the Fort Hood site

mag image **dte** image **cor** image

Figure 2.7 IFSAR components corresponding to the image in Figure 2.6

(a)

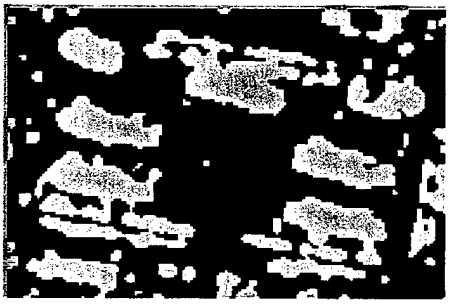




(b)

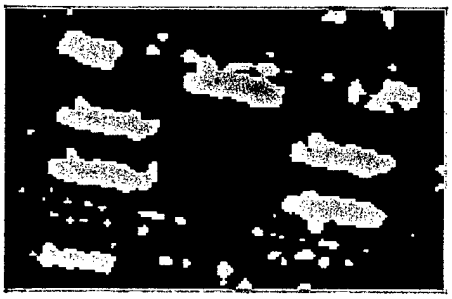
Figure 2.8 Simple thresholding of dte image is not appropriate to derive cues: (a) thresholded at mean elevation; (b) thresholded at mean plus one standard deviation

(a)



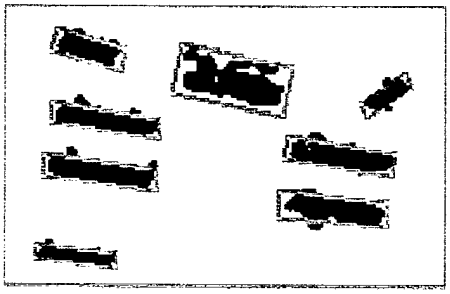
(b)

(c)

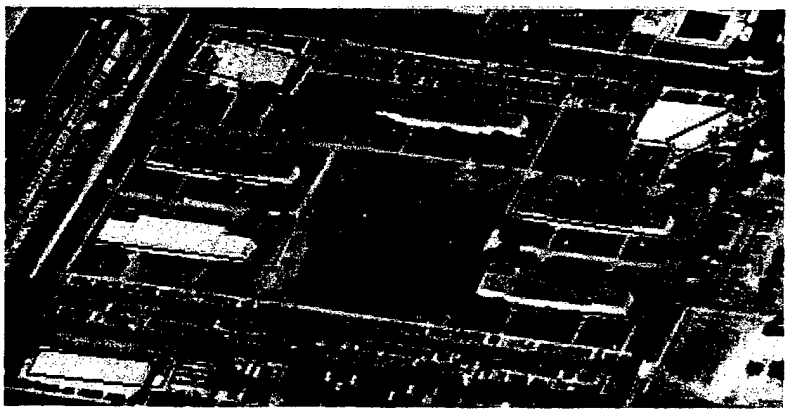


(d)

Figure 2.9 (a) Positive LOG convolution regions from magnitude image; (b) Positive LOG convolution regions from dte image; (e) Combination of (a) and (b) with high correlation (> 0.9); (d) Connected components of (c)



(a)



(b)

Figure 2.10 (a) Approximated bounding rectangles and (b) projected on the panchromatic image

Figure 2.11 SAR mag (left) and IFSAR phase correlation cor images for McKenna MOUT site exhibit characteristics not useful for extraction of building cues

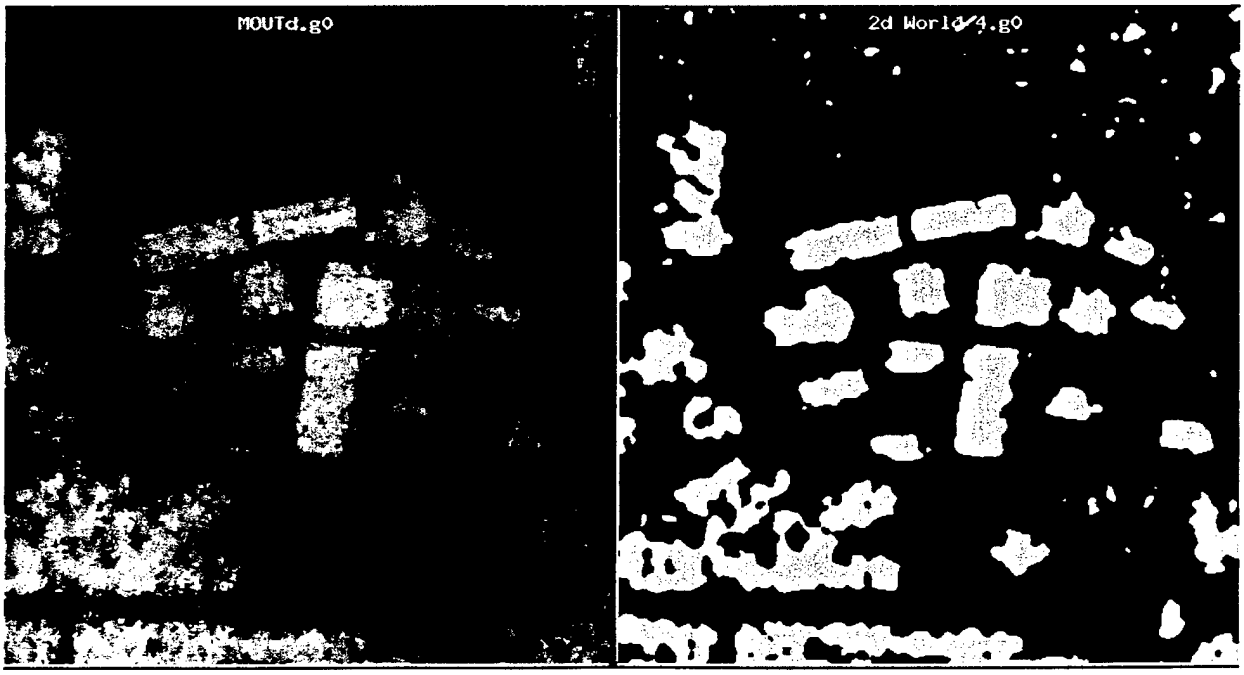


Figure 2.12 IFSAR derived DTE image (left) and 3D object cues extracted (right)

Figure 2.13 Baseline model from SRI International projected on IFSAR cues

3 Description of the Multi-View System

This section describes the multi-image building detection system. As outlined in the introduction section, this system is limited to buildings with rectilinear roofs which project into parallelograms or combinations thereof (assuming locally orthographic projection which is valid for most aerial images). It is assumed that camera models are given, that the ground is flat with known height and that the sun position is given (computable from latitude, longitude and time of day data). Multiple images are *not* assumed to have been taken at the same time.

A block diagram of the system is shown in Figure 3.1. The approach is basically one of hypothesize and verify paradigm. Hypotheses for potential roofs are made from fragmented lower level image features and are verified by using more reliable global evidence. The methodology is to be liberal at the early stages and make decisions only when sufficient information is available to make them reliably.

The system starts by finding lines representing intensity discontinuities in the given images. These lines are then matched across the different views using the constraints given by the known camera geometries and junctions between them computed. These features are then grouped to form the next higher level features which consist of parallel lines or U-structures (representing three sides of a parallelogram), which in turn are used to form parallelogram hypotheses. Simple selection procedures using roof evidence are applied to select among the hypotheses. Wall and shadow evidence is collected to verify the selected hypotheses. Several overlapping hypotheses may result; a choice is made among them by considering all available evidence. The final result consists of 3-D building models. The various stages of this process are described in the remainder of this section.

3.1 Hypotheses Formation

3.1.1 Basic Features

Lines, junctions and parallel lines (called *parallels*) are the basic features used to form roof hypotheses. The higher level process is different for flat roofs than for slanted roofs. The system is hierarchical and uses evidence from all the views in a non-preferential, order-independent way.

Lines

The system starts with line segments detected by linear approximation of edge segments obtained by linking of edges detected by a Canny edge detector. The line segments that are colinear are grouped togethe. r. Segments are considered colinear if there is a *free path* from the end of one segment to the other, i.e. no other segment blocks it, and the angle made by the line joining the closest endpoints of the two segments with each of the segments is less than 10 degrees. Colinearity is applicable to a set of greater than two segments as well; the given criterion must be met between every pair of neighboring segments.

After colinear grouping, the grouped lines are tested for matches across the views by using a quadrilateral constraint. The match for a line segment in one view must lie at least partially within a quadrilateral

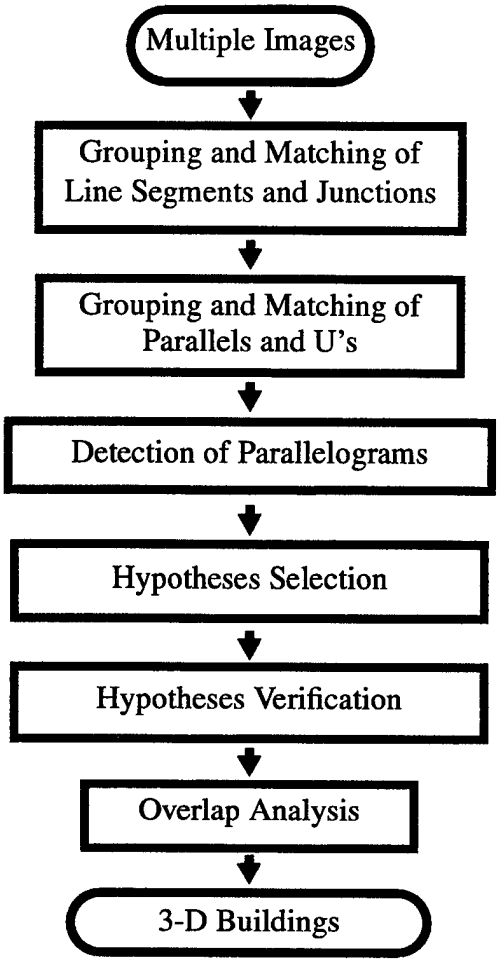


Figure 3.1 Block diagram of the system

defined by the epipolar geometry and constraints on 3-D heights of structures to be detected [Noronha & Nevatia, 1997b].

Each pair of lines that meet the quadrilateral constraint in any pair of views is determined to form a *line match,* and is included in the set of line matches that is denoted by S_{lm} which is used at the higher levels for further processing.

Junctions

Next, junctions are computed from the matched lines in each view. Consider a pair of lines L_{ik} (k = m, n) in $view_i$, with endpoints P_{ikl} (l = 1, 2). Junction J_{imn} is formed at the intersection of L_{ik} (k = m, n) iff the angle between L_{im} and L_{in} is greater than 30^o and min (distance(J_{imn}, P_{ik1}), distance(J_{imn}, P_{ik2})) $\leq$ length (L_{ik}) for (k = m, n). The set of junctions formed in $view_i$ is denoted by S_{J_i}.

Parallels

Next, *parallels*, P_{imn}, are formed between approximately parallel pairs of lines, say L_{im} and L_{in} in $view_i$, that are separated by less than the maximum projected width of a building. Parallels are used as an important feature to form roof hypotheses for both flat and gable roofs. At least one of the two lines forming a parallel must be an element of S_{lm}. If both lines are in the set of line matches and meet some additional criteria, they form *matched parallels*, as defined below.

Matched Parallels

Consider a parallel P_{ik} with component segments L_{ik_1} and L_{ik_2} in $view_i$, and P_{jl} with component segments L_{jl_1} and L_{jl_2} in $view_j$. P_{ik} are said to be matched parallels if exactly one of the following criteria is met:

- (L_{ik_1}, L_{jl_1}) and (L_{ik_2}, L_{jl_2}) are elements of S_{lm}
- (L_{ik_2}, L_{jl_1}) and (L_{ik_1}, L_{jl_2}) are elements of S_{lm}

In the case of parallels over more than two views, the parallel match constraint must be satisfied over parallels from every pair of views. Parallel matches over n views are represented as n-tuples.

3.1.2 Flat Roofs

Flat roof hypotheses are formed from parallelograms. Two different features are used to get parallelograms, which are the matched parallels and the *U* structures (Us represent three sides of a parallelogram as described below). Three-D roof hypotheses are formed from parallelograms by searching for an estimated height. These steps are explained below.

Hypotheses from matched parallels

Matched parallels are used to form parallelograms by finding closures on both sides of them. They are tested for being nearly parallel to the ground. To find closures, the system first looks at matched linears which have similar height to that of the parallel match. If no matched linears of compatible heights are found, then the system uses the longest unmatched linear in the search window. If no linears exist, the system closes the parallel at the ends of the linears forming the parallel. Figure 3.2 illustrates this process.

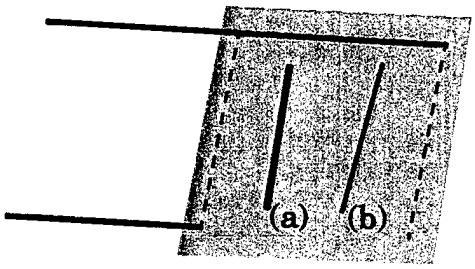


Figure 3.2 Search order for closures; (a) a matched linear with a compatible height; (b) an unmatched linear; (c) the dotted lines (the end points of lines forming the parallel match). The shaded area is the search window

Hypotheses from Us

As a complimentary feature to the matched parallels, U structures are considered. A U captures 3 sides of a parallelogram. A U is formed from a matched line L_{il} with junctions in the set S_{J_i} on at least one side. If the line L_{il} has junctions on both sides, then the three lines L_{il}, L_{im} and L_{in} forming the two junctions J_{ilm} and J_{iln} define the U structure. The side lines L_{im} and L_{in} should be parallel and be on the same side of the seed line. If the line L_{il} has junction J_{ilm} on one side only, a search for an *extended* junction is conducted on the lines which form unmatched parallels with the line L_{im}. Note that only unmatched parallels are used

here as the matched parallels have already been used to form other hypotheses. If a satisfactory line is found on the other side, the three lines again form a U structure. Figure 3.3 shows examples of both case.

Us are computed for each $view_i$, to form sets S_{U_i} (i = 1, 2... number-of-views). The Us are used to form parallelograms by finding closures for the fourth side in the same way as for the case of parallels.

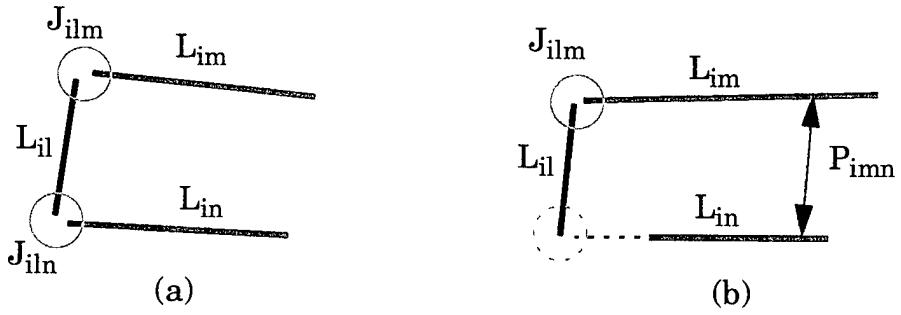


Figure 3.3 U formation from (a) two junctions; (b) one junction and an unmatched parallel

Estimating Roof Height

The hypotheses formed so far are defined in 2-D in each view. Even though the hypotheses may contain matched lines, the matches are not necessarily unique and heights computed from line matches are not reliable. Instead, an estimate for the height of the entire roof is made by conducting a search. The roof is assigned a number of heights in the allowed range in small increments. For each height estimate, the corresponding 3-D hypotheses is projected in each other view and line evidence for each projection is computed as shown in Figure 3.4. The evidence consists of the sum of the lengths of the supporting segments. The height with the best evidence is selected.

Figure 3.5 shows all the hypotheses from the two images of Figure 1.1 and Figure 1.2. Note that there is a large number of hypotheses (1470). In fact, it is not possible to distinguish all of them in this figure as many of them overlap. This large number comes about because the evidence for rooftops is often fragmentary and many nearby structures such as roads, sidewalks, walls, landscaping and other buildings give rise to parallel structures. The philosophy in the design of this system has been that it is preferable to make more hypotheses at the lower levels which can then be filtered at the following stages where more global cues can be considered. Selection and verification of such hypotheses is described in Section 3.3 and Section 3.4.

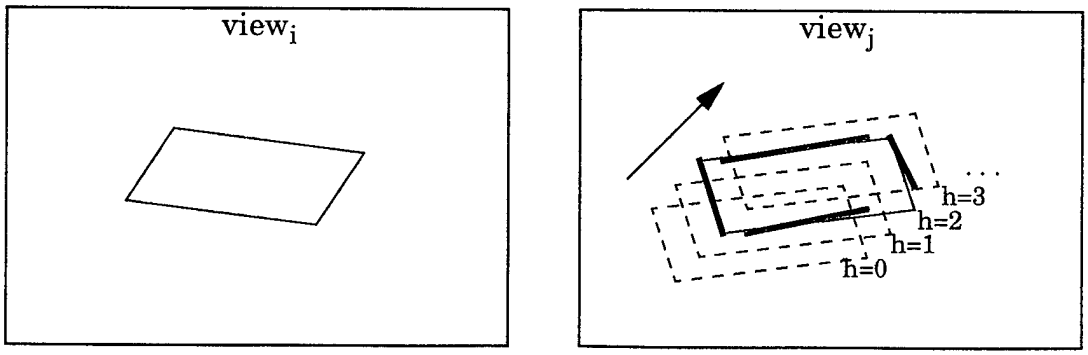


Figure 3.4 Searching the height of a parallelogram

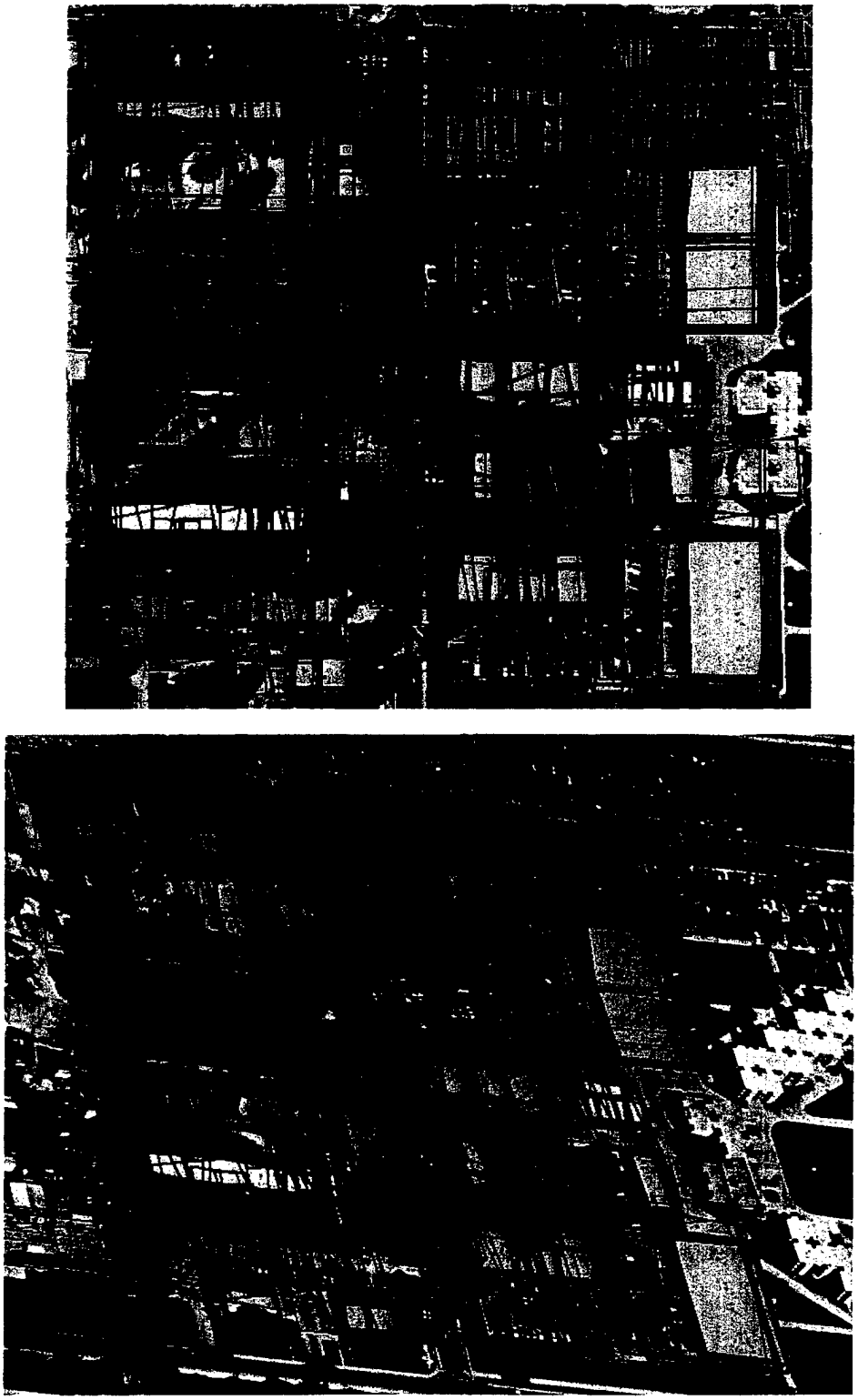


Figure 3.5 All hypotheses from the image pair of Figure 1.1 and Figure 1.2

3.2 Formation of 3-D Hypotheses for Gable Roofs

Buildings with more complex roofs, such as a gable roof, require a more complex hypotheses formation process. This work is still in early stages. The following assumptions are made regarding the model for gable roofs (see Figure 3.6 and Figure 3.7):

- The roof is symmetrical and consists of two rectangular planar surfaces of the same size.
- The overhang is small.
- The two side boundaries and the spine are parallel to each other and parallel to the base boundaries.
- For a given elevation, *h1*, the height of the roof spine varies within the range *a* and the height of the side boundaries vary in the range *b*. The width of the roof is within a minimum and maximum width.

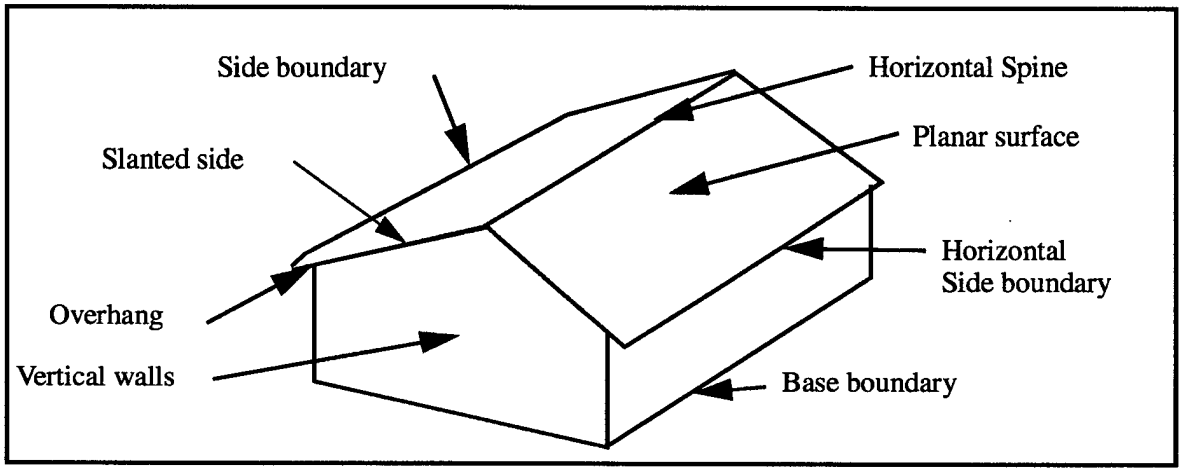


Figure 3.6 Model of building with a gable roof

The hypotheses formation process consists of two parts. The first is a 2-D process that forms the seeds for the hypotheses from features extracted from each view (though it does use matching information from the other views). The second part uses support information from the other views to derive the 3-D information needed to construct the model. The details are as follows (see Figure 3.8 and Figure 3.9)

2-D Process (for each view):

- Use the matched lines having a pre-determined minimum length among the various views as input (see example below in Figure 3.10.)
- Form all parallels among these lines that are within a specified range of distances in 2-D.
- For each parallel formed in the previous step, search on both sides, for a third parallel to form a triple of parallel lines (see Figure 3.7)

3-D process (for each view):

- For each triple of parallel lines, take the middle line (gable spine) and collect its matches in a second view, that lie in the range *a* of heights (Figure 3.8 (a)).
- For each match for the gable spine, collect the matches of the side parallels of the triple that lie within the range *b* of heights, with the additional constraint that the sides can not be at a higher elevation than that of the spine of the roof (Figure 3.8 (b)).
- Form 3-D triples from the combinations of matched spine and side boundaries.

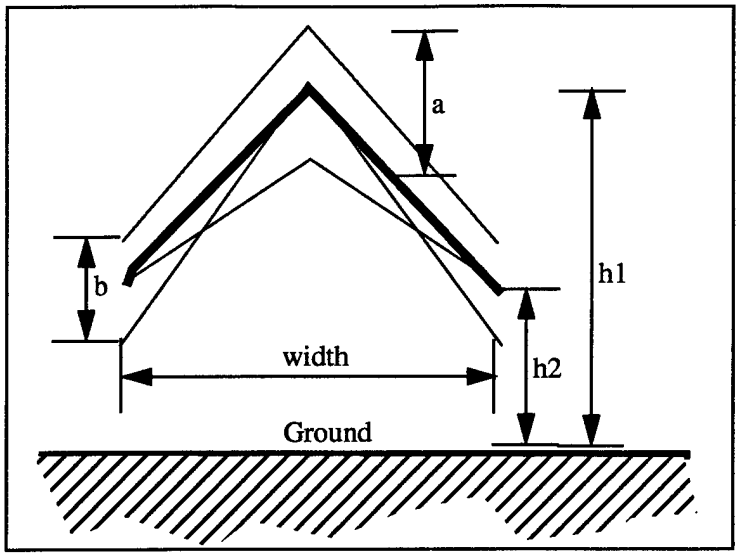


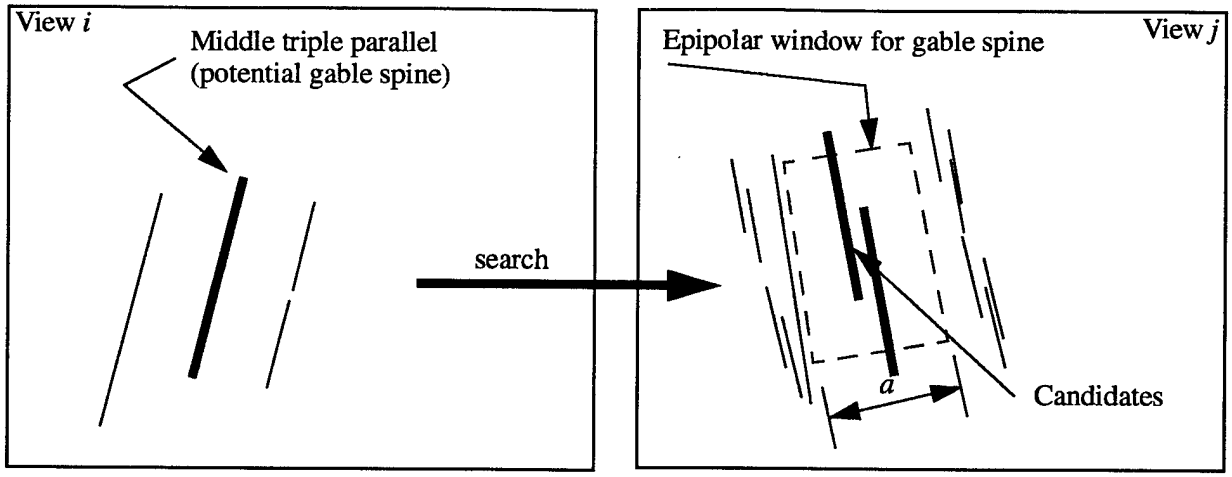
Figure 3.7 Symmetric gable roof degrees of freedom

- Use the triple to determine closure for 3-D roof surfaces, as shown in Figure 3.9. We calculate the slanted side boundaries by intersecting the ends of the spine and the two side boundaries. The constraint is that the slanted sides must form 90-degree junctions in 3-D with the spine line and with the side boundaries.

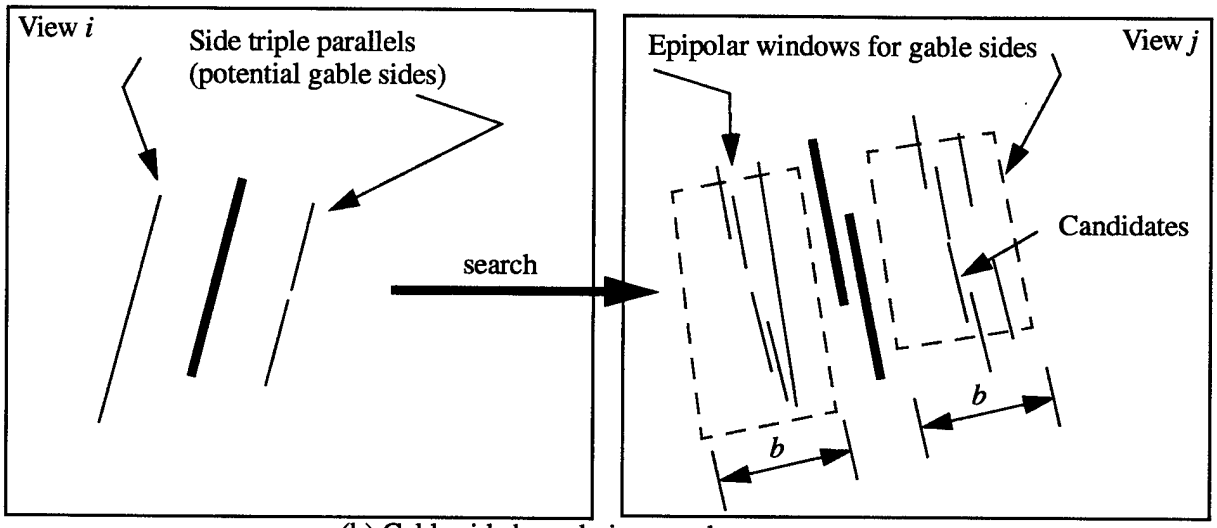
The resulting set of hypotheses is constructed from the union of the sets of hypotheses computed from each view.

As an example consider the two windows shown in Figure 3.10 corresponding to a portion of the McKenna MOUT site at Fort Benning, GA. The figure shows the input lines to the hypotheses formation process. The set of hypotheses formed from these two views are shown in Figure 3.11 overlaid on both views. The parameters used for this result are as follows:

- The minimum length of a matched line to consider in forming triple parallels: 6 meters.
- The minimum separation between triple parallel elements: 5 meters.
- Width of height layer for gable spines: between 8 and 10 meters.
- Width of height layer for gable sides: between 2 meters and the height of the spine.
- Minimum height of gable sides: 3 meters.
- Maximum width of building: 10 meters.
- Maximum length of building: no limit.



(a) Gable spine search process



(b) Gable side boundaries search process

Figure 3.8 Search for roof elements to calculate 3-D position of roof hypothesis

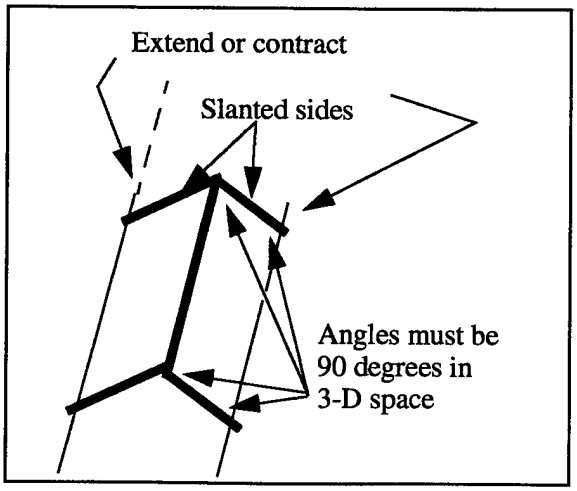


Figure 3.9 Closure of 3-D gable roof hypotheses

Figure 3.10 Matched lines (two views)

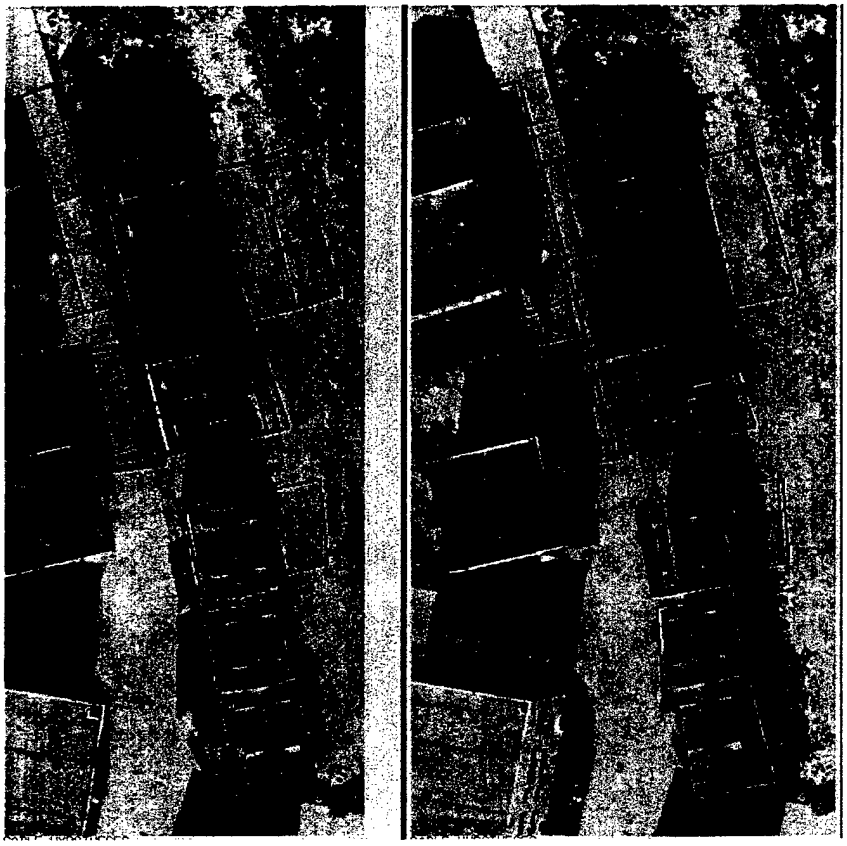


Figure 3.11 Gable hypotheses formed from two views.

3.3 Hypotheses Selection

As indicated earlier, the system makes a large number of hypotheses compared to the actual number of buildings present in a scene. Good hypotheses can now be chosen based on additional evidence. It is computationally expensive to collect all the available evidence for each hypotheses. To reduce this complexity, a *selection* process using a sequence of filtering steps using relatively inexpensive evidence is applied first; it is able to reduce the number of hypotheses to be considered further significantly. After selection, more evidence that is computationally expensive to collect is applied in the *verification* stage. The selection process is described below.

3.3.1 Hypotheses Selection for Flat Roofs

Roof Line Analysis

Roof line evidence is one of the less expensive evidence to obtain. It is obtained by a search performed in each view for evidence supporting, or negating, a roof hypothesis. Lines that are found within a certain distance (which is a function of the image resolution) from the hypothesized parallelogram, and that differ in angle by not more than a certain angle, are considered positive evidence. Negative evidence consists of lines that cross the boundaries of the hypothesis. Figure 3.12 illustrates the concept of positive and negative evidence. The strengths of the positive and negative evidence depend on the lengths of the image lines supporting them. Hypotheses with low positive evidence or high negative evidence are filtered out.

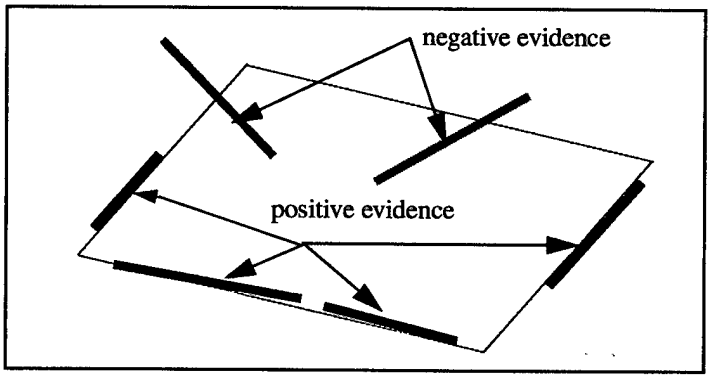


Figure 3.12 Positive and negative line evidence

Using IFSAR Cues

If pre-calculated IFSAR cues are available, they are also used in this stage. For each hypotheses, support from IFSAR analysis is calculated. The hypotheses is projected onto the IFSAR image and overlap of the projected roof with IFSAR regions is computed. The current system requires that the overlap be at least 50% of the projected roof area.

Overlap Analysis

Many overlapping hypotheses are typically formed for the same areas in a scene. This can be due to multiple nearby hypotheses formed from nearby lines in a single view or by similar hypotheses being formed in each view independently. To reduce the effort required in the verification stage, a coarse overlap analysis is performed as the final filter in the selection process. For each set of *significantly* overlapping hypotheses, the roof line support, consisting of the difference between the positive and negative evidence as defined earlier, is compared and the best half are chosen. Two hypotheses are considered to overlap *significantly* if more than the half of the areas of the regions overlap.

Figure 3.13 shows the selected hypotheses from those shown earlier in Figure 1.1 and Figure 1.2. In this example, only 176 hypotheses are selected from 1470 hypotheses.

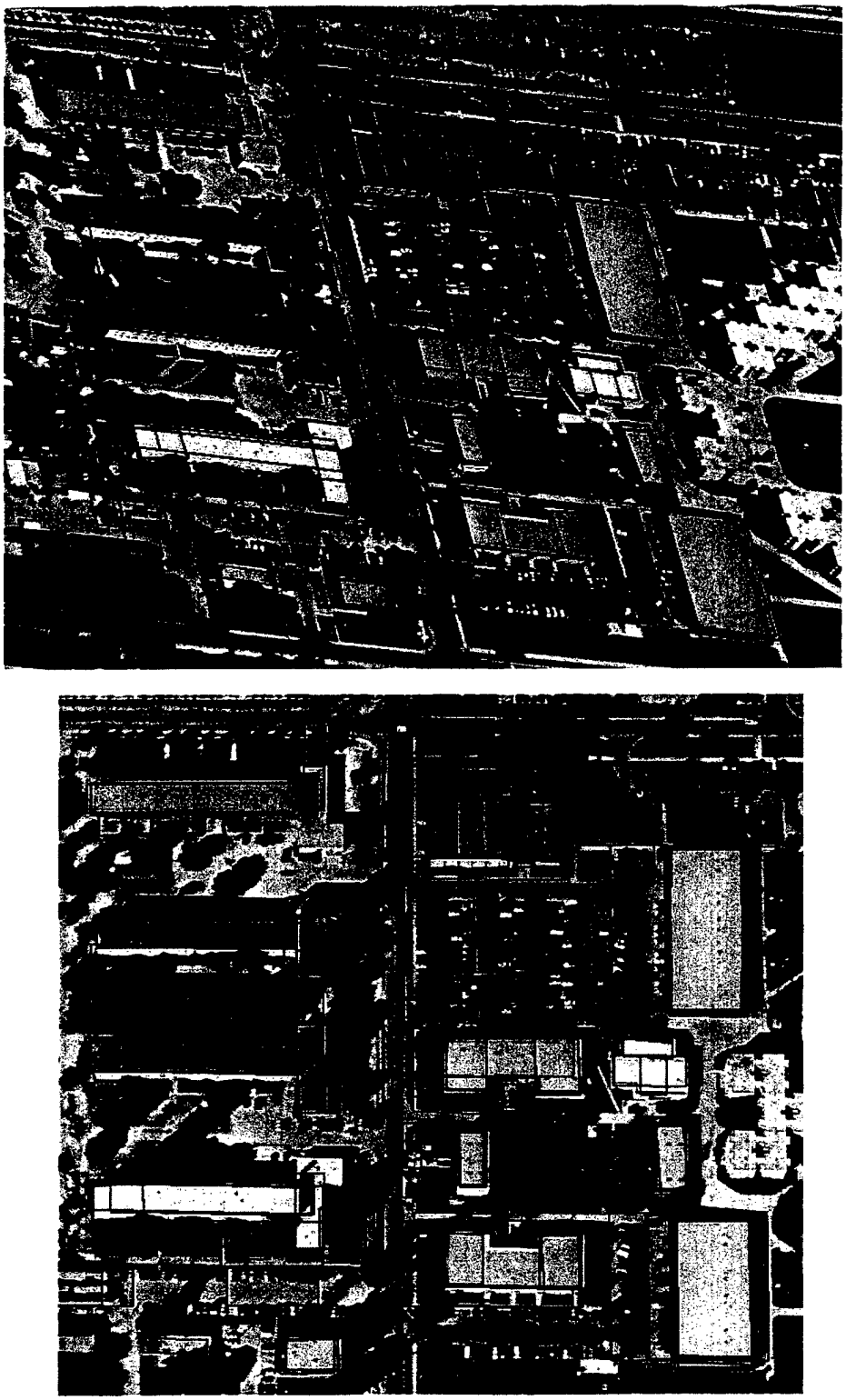


Figure 3.13 The selected hypotheses from the images of Figure 1.1 and Figure 1.2

3.3.2 Hypotheses Selection for Gable Roofs

Hypotheses selection for gable roofs should follow steps similar to those for flat roofs. However, the various procedures needed to collect the required evidence have not been completely implemented yet. At the current time, the only selection performed for gable hypotheses is by using the cues derived from the IFSAR analysis.

Figure 3.14 show the portion of the "cues" image shown earlier in Chapter 2 that corresponds to the image portion processed. The gable roof hypotheses are shown overlaid on the cues image. We require that a minimum of 60% overlap exists in the intersection of each hypotheses foot print and the cues represented in the cues image. The hypotheses selected on this basis are shown in Figure 3.15, from those shown earlier in Figure 3.11. These process efficiently helps filter out hypotheses that clearly do not correspond to roofs or other elevated objects.

Figure 3.14 Testing gable roof hypotheses against IFSAR cues

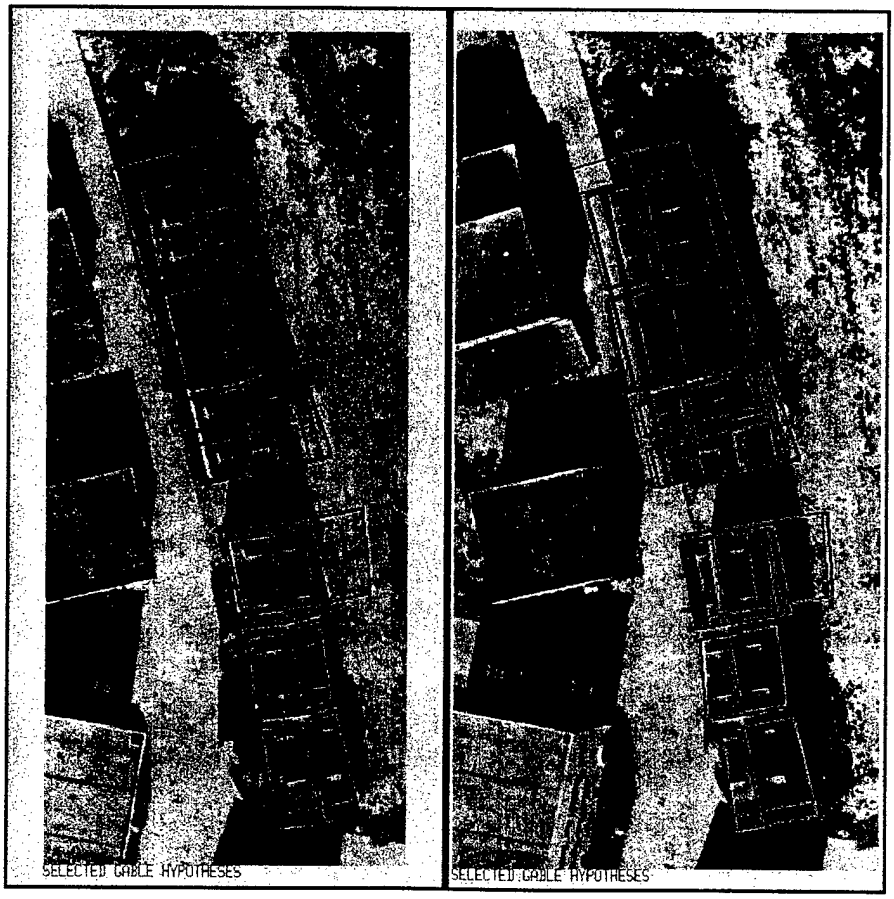


Figure 3.15 Hypotheses selected using IFSAR cueing

3.4 Hypotheses Verification

The next step is to verify whether the selected hypotheses have additional evidence for corresponding to being buildings. This evidence is collected from the roof, the walls and the shadows that should be cast by the building.

3.4.1 Evidence Collection

Figure 3.16 and Figure 3.17 show the wall and shadow elements involved in collecting evidence for verification.

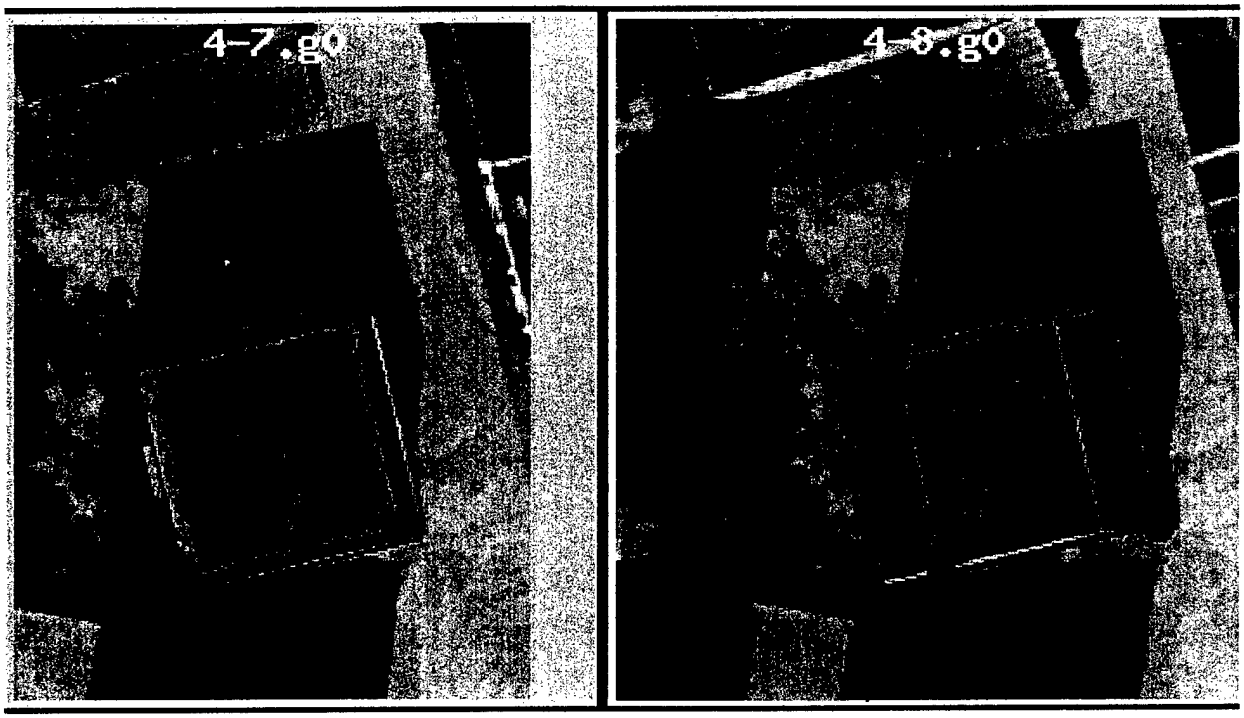


Figure 3.16 Wall elements: roof outline (green), visible base and vertical boundaries (cyan), non-visible elements (red)

Since the hypotheses are represented in 3-D, deriving the projections of the walls and shadows cast, and determining which of these elements are visible from the particular view point is possible. These in turn guide the search procedures that look in the various images for evidence of these elements among the features extracted from the image.

Roof Evidence

Roof line evidence has already been collected in the selection stage. Roof region evidence is also collected now. Since most of the building roofs have no texture in them, the intensity variation of the roof area is small and useful as a roof region evidence. The standard deviation of the intensity is calculated and used.

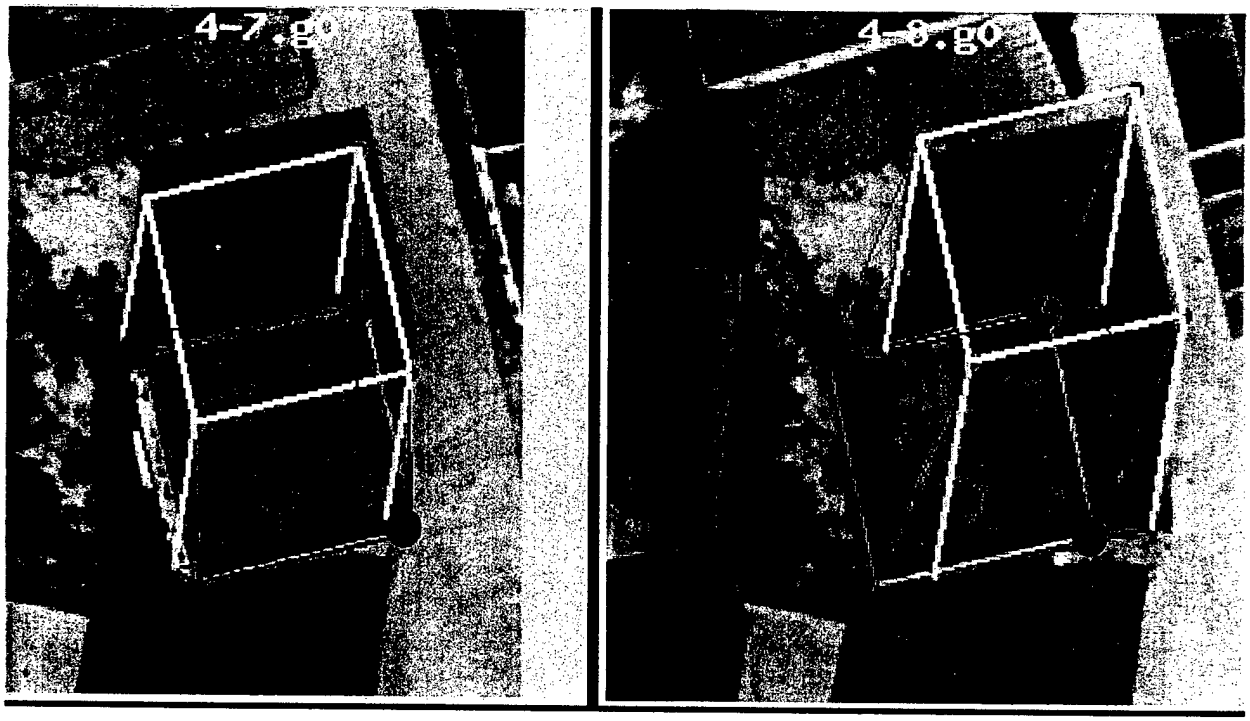


Figure 3.17 Shadow elements: Cast by roof outline (green) on the ground (yellow); cast by vertical boundaries (magenta) on the ground (brown)

Wall Evidence

We look among the line segments extracted form the images for evidence of vertical edges and for evidence of boundaries corresponding to the base of the buildings. Figure 3.18 a shows, in cyan color, the line segments found near the base of the building to support the hypotheses shown in yellow.

A score is computed for each evidence element. For the "cube" object in the example, two vertical edges and two base edges are expected to be visible. For each one of these, the percentage of the length of the found elements compared to the lengths predicted by the elements of the 3-D hypotheses are computed. Separate scores are computed for the vertical and for the non-vertical elements.

Shadow Evidence

Shadow evidence is computed by projecting the roof outlines onto the ground surface. It is assumed that the terrain immediately adjacent to the buildings is flat and level and that the shadows are cast on the ground. In some cases these assumptions are only partially valid and as a result only partial evidence may be found. Since the most important evidence comes from the correspondence of roof elements across views, the added complexity introduced by more sophisticated shadow analysis (shadows cast on irregular terrain or on other structures) is not justified.

Evidence of shadows cast by vertical wall boundaries, by horizontal boundaries, and for the junctions formed by these is searched for. Here the search space is reduced to those line and junctions in the various image views that are potential shadows. These are determined by comparing the projection of the sun rays on the image and the orientation of the line segments. The chosen line segments, shown in Figure 3.19 correspond two groups of shadows: 1) those oriented parallel to the projection of the sun rays on the ground,

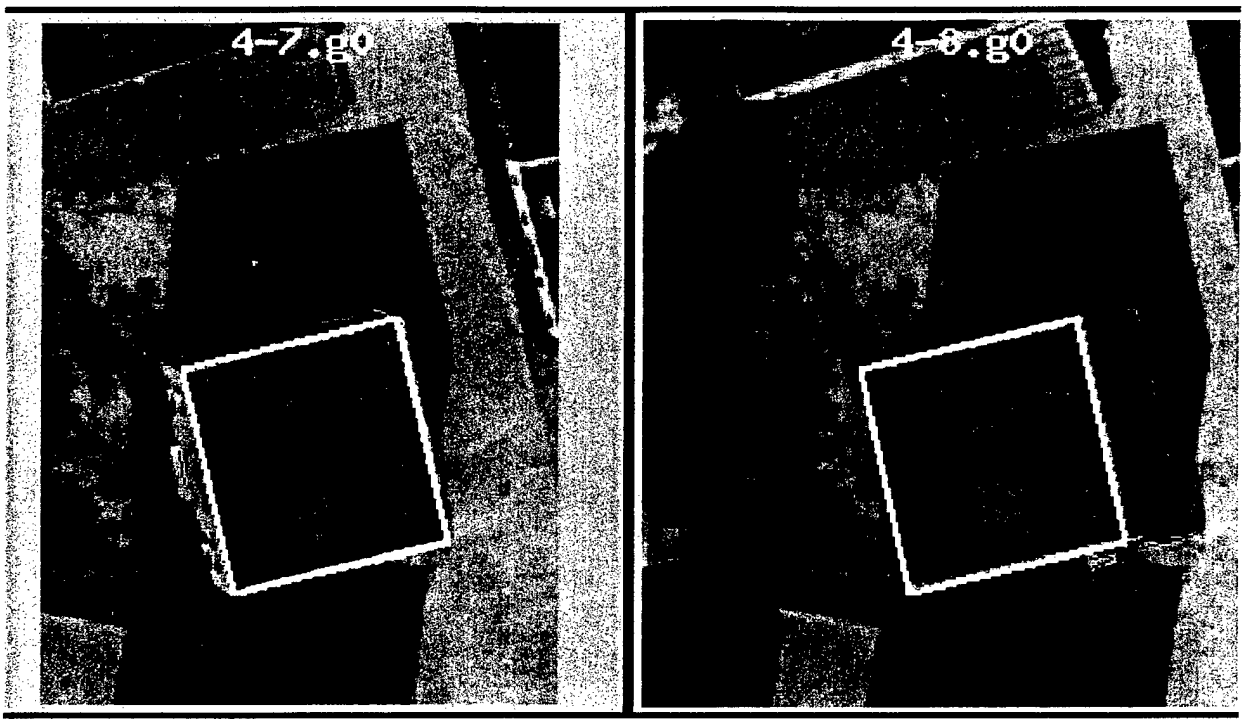


Figure 3.18 Evidence of wall elements found (cyan) in the image

potentially cast by vertical object edges and 2) those potentially cast by horizontal object edges. The darker region on either side of the line segments must be consistent with the direction of illumination.

Shadow evidence is collected also for the junctions formed by the roof and side edges of a building. Junctions are strongly localized features and represent strong evidence. "Strong" object junctions (shown by a green circle in Figure 3.17) correspond to junctions formed by two shadow casting roof edges. Their projection on the ground determines where to look for the corresponding shadow junctions in the image. "Weak" junctions correspond to junctions formed by one horizontal and one vertical building edge (shown by cyan circles in Figure 3.17.)

Separate scores are computed for the different shadow elements. The presence of matching junctions is denoted by 1 or 0. The linear elements (cast by verticals or by non-verticals) contributions are calculates as percentages. Shadow region average intensity is also computed to see that it is sufficiently "dark", this score is a 1 or 0.

The shadow line and strong junction evidence found for the building in our example is shown in Figure 3.20. The roof outline is shown in green. The evidence of shadows cast by verticals is shown in yellow. The remaining shadow boundaries found are shown in cyan.

3.4.2 Verification

Each of the collected evidence composed of smaller pieces of evidence. A critical question is how to combine these small pieces of evidence to decide whether a building is present or not and how much confidence should be put in it. A variety of methods for this are available such as linear weighted sums of components, decision trees, certainty theory, neural networks and statistical classifiers. In the past, such decisions have been made largely on an *ad hoc* basis and the parameters associated with them have been

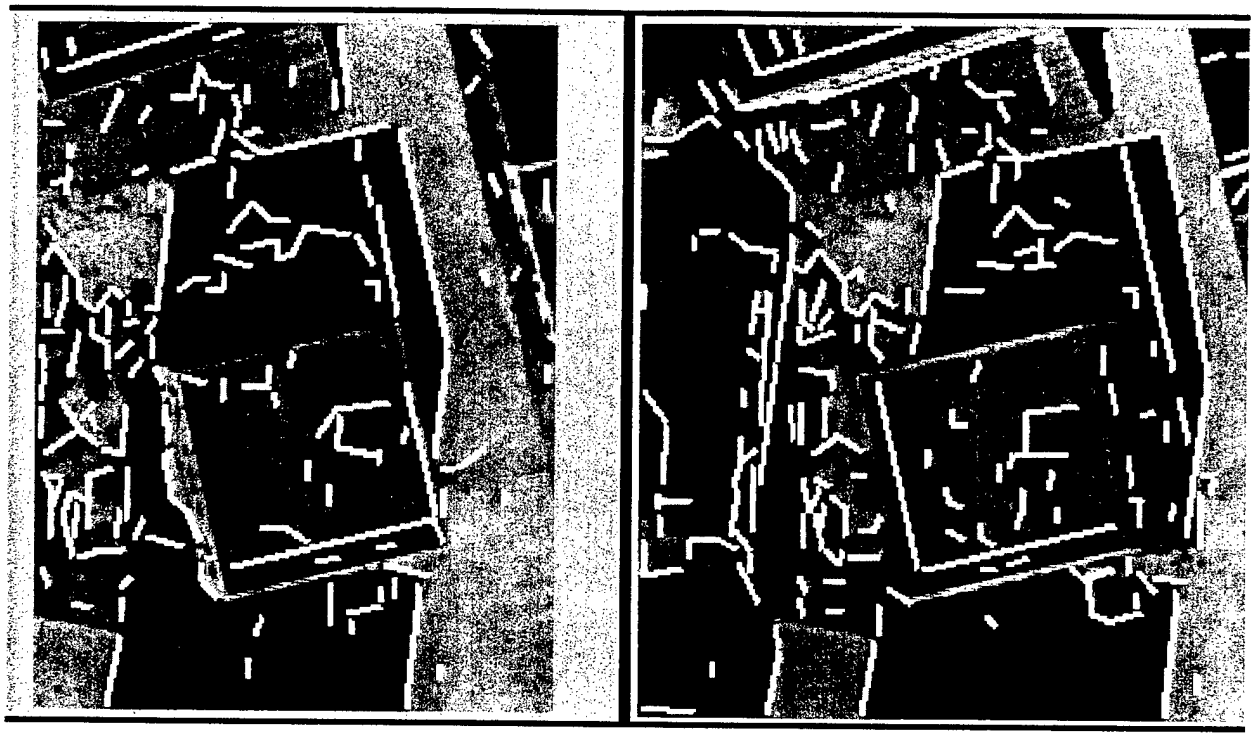


Figure 3.19 Shadow boundaries search space of potential shadows. Junctions among these are also computed and searched for

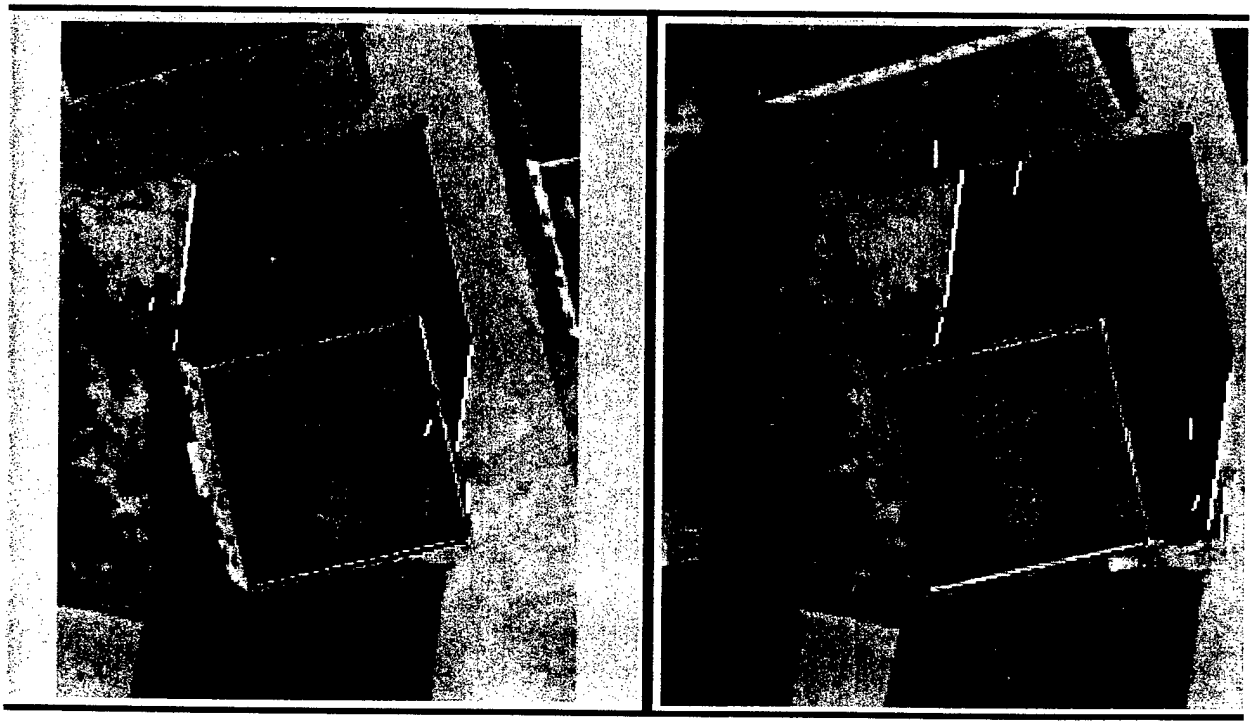


Figure 3.20 Line and strong junction shadow evidence found for this roof hypothesis

selected by the developer testing on a small set of examples. Such approaches suffer from possible lack of generalization. In the current work, attempt is being made to use a Bayesian classification approach which is optimal if some underlying assumptions are satisfied. Further, it is possible to use machine *learning* approaches to automatically estimate the parameters needed for the classifier. Basic research in use of Bayesian approach and learning procedures at USC is being supported by a grant from the Army Research Office (grant # 530-1416-01), the multi-image system described here is serving as a testbed for this methodology.

Bayes' rule provides the probability of a hypothesis being a building given a set of evidence $\boldsymbol{E}$ as follows:[1]

$$P(B|\boldsymbol{E}) = \frac{P(B)P(\boldsymbol{E}|B)}{P(B)P(\boldsymbol{E}|B) + P(\neg B)P(\boldsymbol{E}|\neg B)}, \qquad \text{(EQ 1)}$$

where $P(\boldsymbol{E}|B)$ is the probability distribution of $\boldsymbol{E}$ given that the hypothesis is a building, $P(\boldsymbol{E}|\neg B)$ is the probability distribution of $\boldsymbol{E}$ given that the hypothesis is not a building and $P(\boldsymbol{B})$ is the prior probability of a hypothesis being or not being a building. The difficulty in applying this rule is that the evidence $\boldsymbol{E}$ consists of several components say $\boldsymbol{E}_i$ and hence $P(\boldsymbol{E}|B)$ and $P(\boldsymbol{E}|\neg B)$ are joint distributions of a high dimension. For the problem of building verification, this joint distribution. is not available from mathematical analysis and is difficult to estimate empirically due to its high dimensionality.

To simplify the process, a gross assumption is made that the various pieces of evidence are statistically independent, in which case $P(\boldsymbol{E}|B) = \prod P(\boldsymbol{E}_i|B)$ for all instances $\boldsymbol{E}$ of $\boldsymbol{E}$. A classifier with this assumption is sometimes called a *naive* Bayesian classifier [John & Langley, 1995]. The probability distributions for the naive Bayesian classifier are obtained simply by histogramming the observed values of $\boldsymbol{E}_i$ for a number of examples of hypotheses that represent buildings and not buildings. More accurate approximations of the joint p.d.f. can be made by allowing dependence of some of the parameters in a multi-layer Bayesian *network*, however, estimating parameters of this network is much more complex than for the naive Bayesian classifier; this approach will be explored in future work.

The Bayes classifier is applied to the *selected* hypotheses to compute the probability that they correspond to a building based on their associated roof, wall and shadow evidence. Hypotheses with probability lower than a threshold are discarded, the remaining ones are called *verified* hypotheses. Figure 3.21 shows the verified hypotheses on the headquaters area. The confidence values are shown by color coding: red indicates very high probability (0.85~1.0), orange indicates medium probability (0.65~0.85) and yellow indicates low probability (0.5~0.65).

3.4.3 Overlap Analysis

At this stage of processing, several overlapping verified hypotheses may remain as can be seen in Figure 3.21. Only one hypotheses of the *significantly* overlapping hypotheses is selected. For the overlap analysis, the evidence is recomputed since the this stage requires finer analysis than the verification process. Therefore, the wall and the shadow evidence is recomputed with narrower search regions.

Overlap analysis is done by one-on-one comparison of the overlapping hypotheses. For this analysis, another naive Bayesian classifier is applied to compare two hypotheses. The inputs to the classifier are the differences between collected evidence of the two hypotheses. The output is a probability of the first hypothesis being closer to the actual building.

---

1. Bold letters represent random variables while plain letters are probability instances.

The verified hypotheses for Figure 1.1 and Figure 1.2 are shown in Figure 3.21 and the final results are on Figure 3.22. The results show that most buildings are detected, either completely, or in part. The bright building at the bottom left is not detected as the hypotheses corresponding to it is not verified; it has poor roof line evidence as the lower boundary of the roof is not visible in one view and has a low contrast in the other (this building is detected if a different pair of views is used as shown later in Chapter 5).

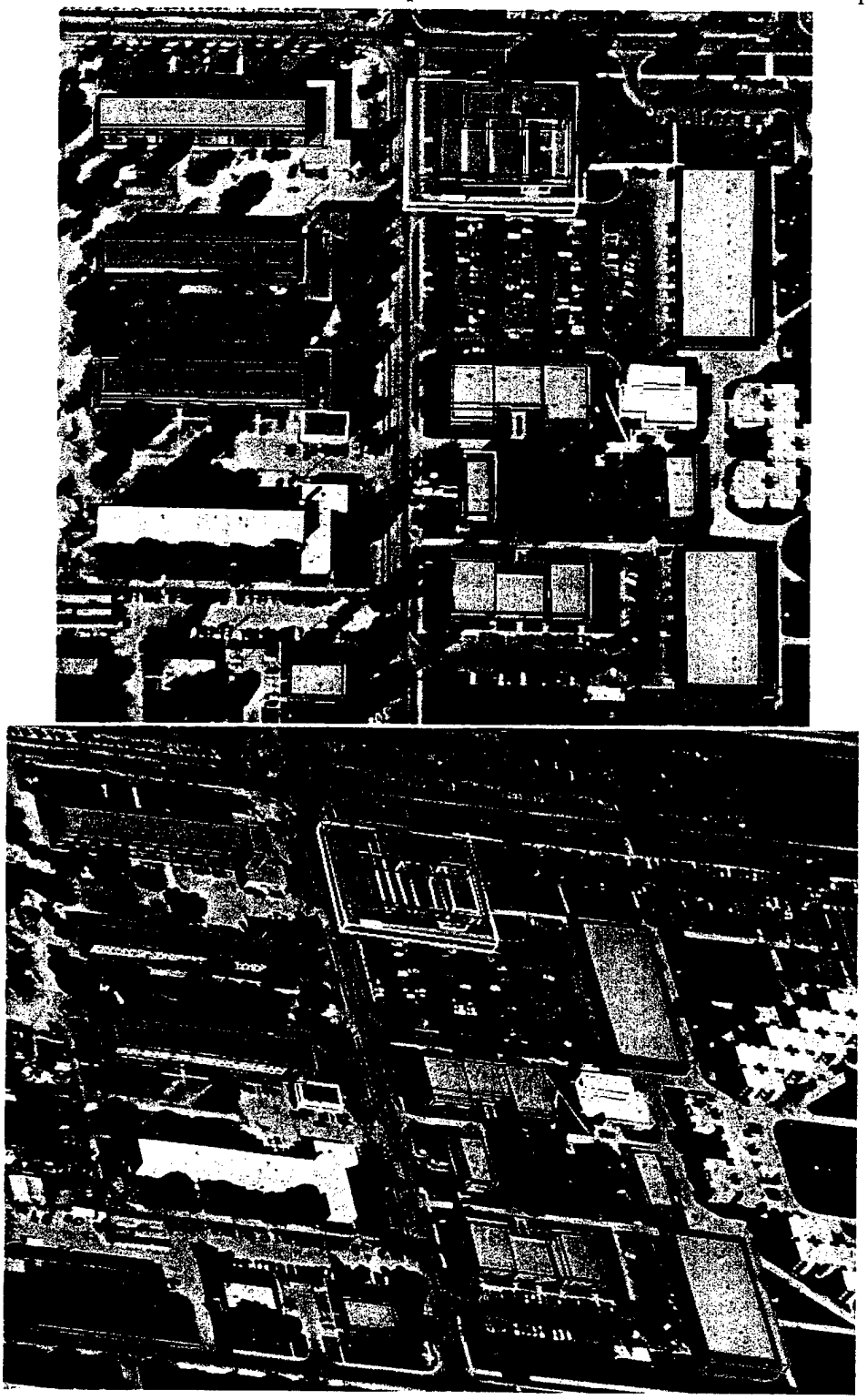


Figure 3.21 Verified hypotheses from the image pair of Figure 1.1 and Figure 1.2

Small parts of the complex buildings are not found, they are also rather low structures. Buildings with concavities are not delineated completely accurately. There are only two "false alarms" (*i.e.* buildings found where there are none). One of these is of low confidence (yellow) and could be easily eliminated by raising the acceptance threshold. More examples and discussion of results is given in Chapter 5.

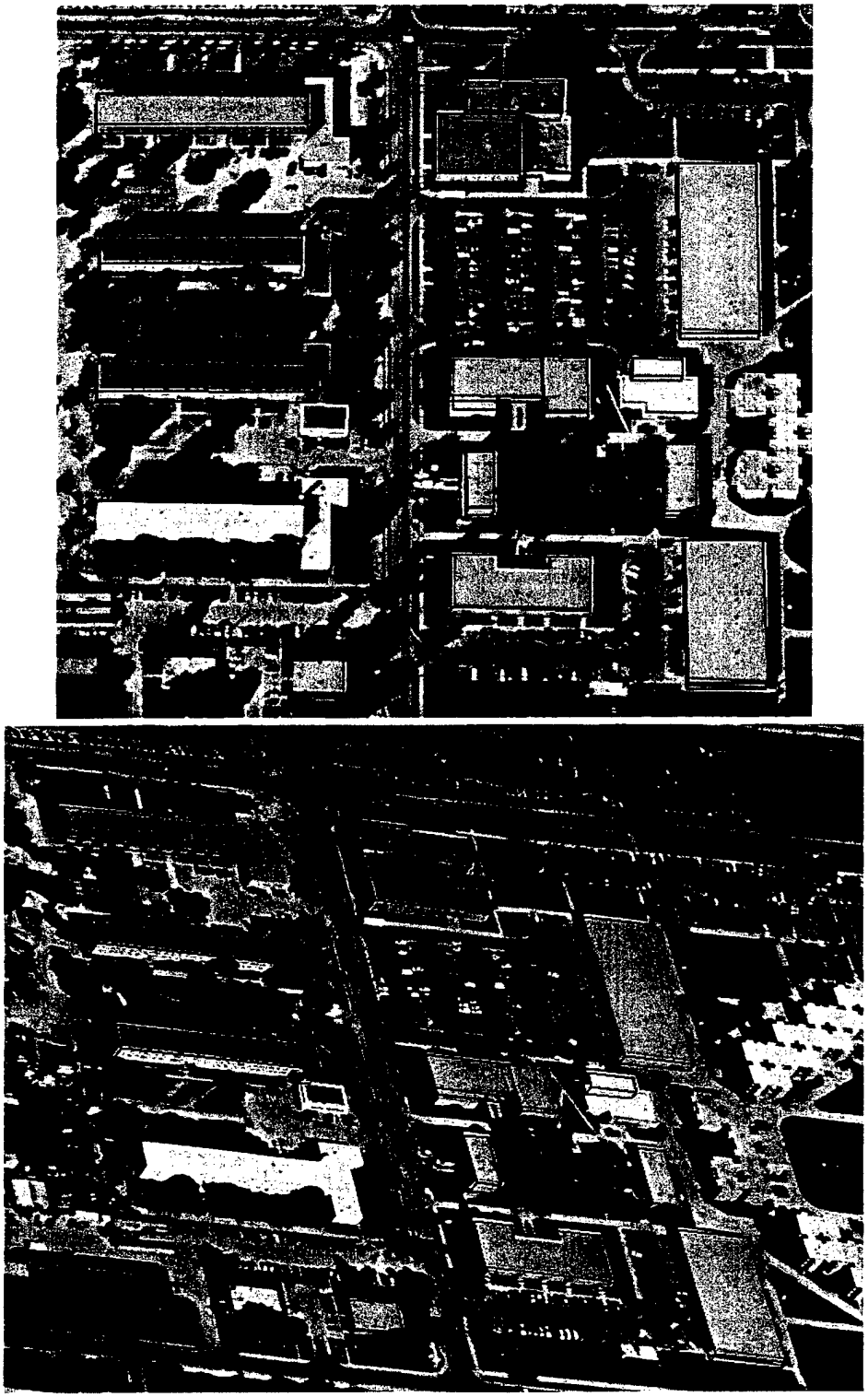


Figure 3.22 Final results from the images of Figure 1.1 and Figure 1.2

4 Interactive Corrections

While the performance of the described automatic systems believed to be an advance over previously available techniques, the building models need to be refined further to meet the needs of most application tasks. For this purpose, an interactive methodology has been developed that can use partial results of the automated analysis. This method follows the ones described in [Heuel & Nevatia, 1995; Noronha & Nevatia, 1998], but is made easier due to availability of multiple images.

The underlying approach for the development of the interactive system is that it must be efficient in all aspects; it must minimize the need for user inputs and it should have an interface that provides simple user interaction. The current method is designed for rectilinear buildings with flat roofs. The user interaction may occur either after a complete run of the automatic system or only after the stages where lines have been matched and junctions detected. User interactions aid the automatic system in forming new hypotheses. Three-D height computations are still performed automatically. The system requires a user to interact in one image only, even though a second view is displayed and used by the automatic system. This can greatly reduce the effort required of the user. Building parts can typically be detected by just a few, imprecise mouse interactions (called *clicks* from here on).

The operation of the interactive corrections is described below. Two situations are distinguished. In the first situation a building has not been detected and needs to be added. In the second, a building has been detected partially and requires editing. For the purpose of the discussion a "click" indicates a click of a button on a mouse or other pointing device.

4.1 User Interactions

4.1.1 Adding a Building

A new building can be added by typically a maximum of **three** clicks to generate an accurate 3-D representation of a rectangular component. In many cases a **single** click is sufficient. Each click consists of the user pointing to a corner of the building; the pointing need not be precise as the precise corners are selected automatically from the image data. Before the start of this session, line matches and junctions need to have been computed by the automatic system.

The system operation is as follows:

1.The user clicks on or near a corner on the roof of the desired building.

2.The system attempts to form a building hypotheses from this information and displays the results to the user. If the results are satisfactory, user can indicate so, else

3.The user clicks on a second corner of the roof. The system repeats its attempt to form a hypotheses as before and displays the results. If the results is not satisfactory,

5.The user clicks on a third corner on the roof. Three corners suffice to define a parallelogram for the roof; height is computed automatically and may require an extra correction in rare cases.

4.1.2 Editing a Building

This process can be used to edit a current building hypotheses, either derived automatically or by interactions in earlier stages as described above. The three available corrections are to adjust a corner, adjust a side or to adjust height. If more complex interactions are needed, the building should be deleted and reconstructed. Again, interactions take place in only one view and only approximate locations are needed. The user interactions are as follows:

- To adjust a corner, click on a new location for it.
- To adjust a roof side, click anywhere along the actual roof side.
- To adjust height, click anywhere along the base of the building.

For each of the steps above, the system recomputes all aspects associated with the formation of 3-D hypotheses during automated operation. These are described in the next section.

4.2 Automated System Tasks

This section describes the action of the automated system as a response to the user clicks discussed above.

4.2.1 Adding a Building

Figure 4.1 depicts the situation after the first user click.

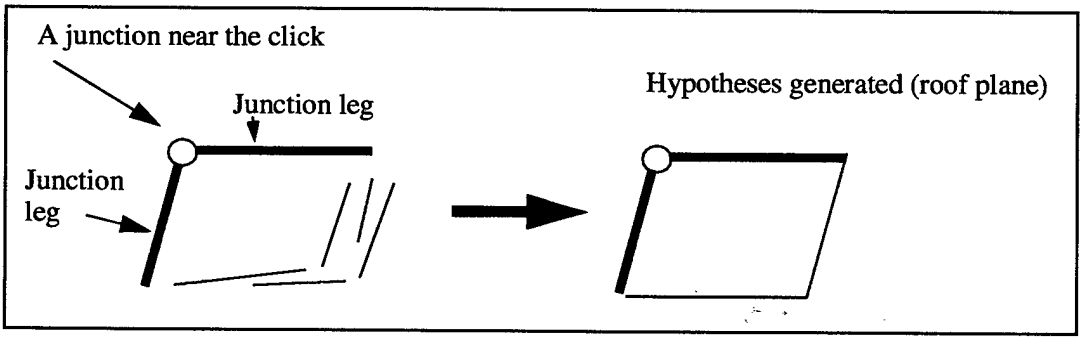


Figure 4.1 First input analysis

The system locates all junctions near the click and reports failure if none is found. For each junction found, the system attempts to construct a parallelogram. The parallelogram is formed by first examining the stored information looking for a U-structure that uses the junction legs. If no U-structure is available, the junction legs are used to derive the parallelogram (roof hypothesis). The elements of the parallelogram are matched to elements on the other views and scores are computed as the system would during automatic operation. The system then selects one configuration and presents it to the user.

Figure 4.2 illustrates the situation after the second click. The second click is used to generate a new hypotheses in the same manner as with the first click. The hypotheses are formed that include the point from the first click however, are weighted higher.

After the third click, the three points are used to form three possible parallelograms to represent roof hypotheses, as shown in Figure 4.3. The system calculates the 3-D orientation of these planes, matches the elements with elements on other views. For all possible matches, select those hypotheses with least inclination for a flat roofed building. Also the angles between the sides must be close to 90 degrees in 3-D. The system computes scores as before, and selects the hypothesis with the best score.

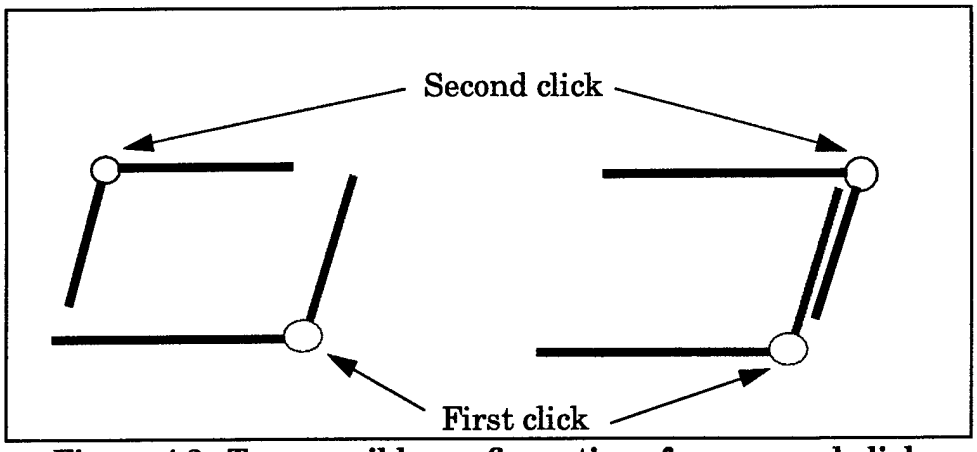


Figure 4.2 Two possible configurations for a second click

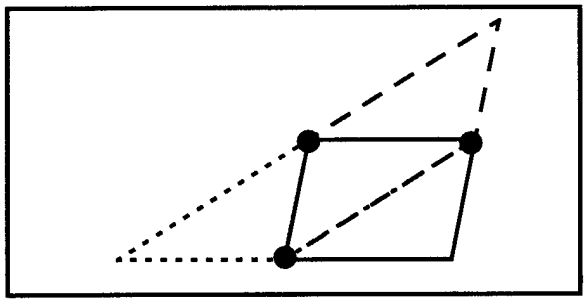


Figure 4.3 Three parallelograms can be formed from three points

Next we show some examples of user interaction using portions of images from the McKenna MOUT site at Fort Benning. Figure 4.4 shows two examples where a single click was sufficient to recover each of the two buildings, with no further editing required.

Figure 4.5 shows an example where three clicks are needed. The first click results in a partial hypotheses. The second click is not sufficient to obtain the correct hypotheses. The third clock results in an accurate model that requires no further editing.

4.2.2 Editing a Building

The actions of the automatic system for editing a building are similar to those for adding a buildings. If a corner is indicated by the user, the system finds the nearest corner in the existing hypotheses and replaces it by a corner near the indicated position. A new hypotheses is generated and its height recomputed. If a side is indicated, the closest side of the existing hypothesis is found and moved to include the indicated position. Again, a new hypothesis and building model is constructed. The height correction is a little more complex. Basically, the automatic system looks for a height such that one of the projected wall bases now passes through a point indicated by the user. These processes are illustrated by examples below.

Figure 4.6 shows a detected building that is only partially correct. The appropriate correction consists of indicating a point along the actual building boundary to cause the system to adjust the incorrect side. As before, the system automatically recalculates the height and location of the new model.

The last example, shown in Figure 4.7 illustrates a similar procedure to adjust the height of the building. The user needs only to select a point along the base of the building.

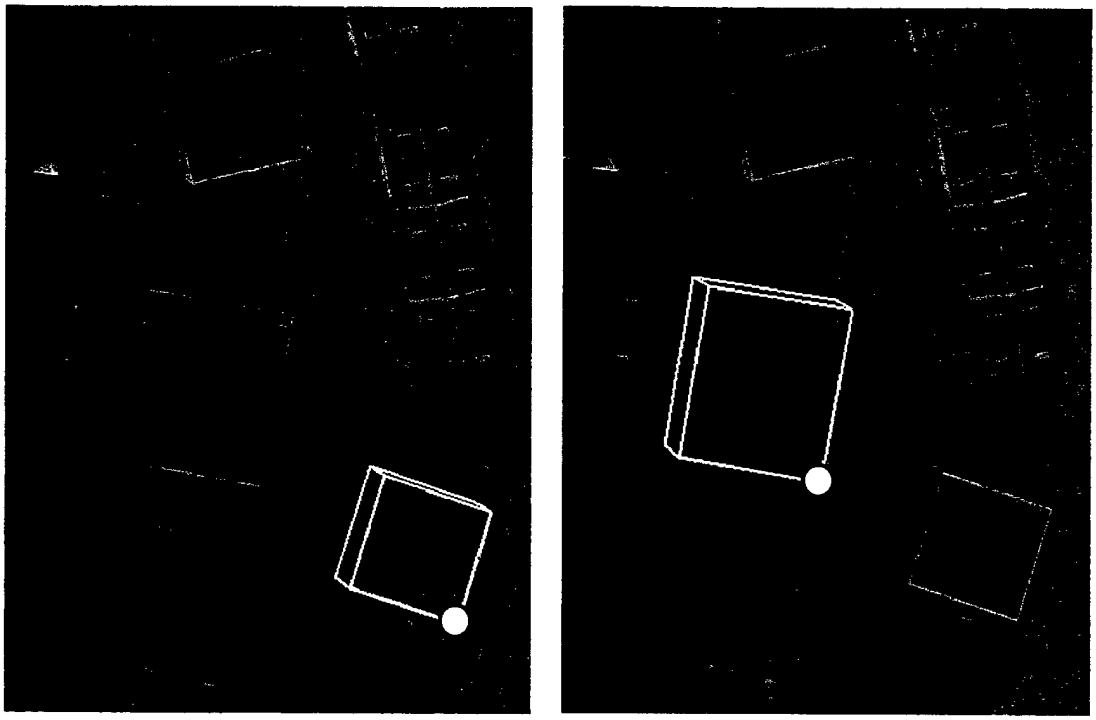


Figure 4.4 Addition of these buildings required one click each

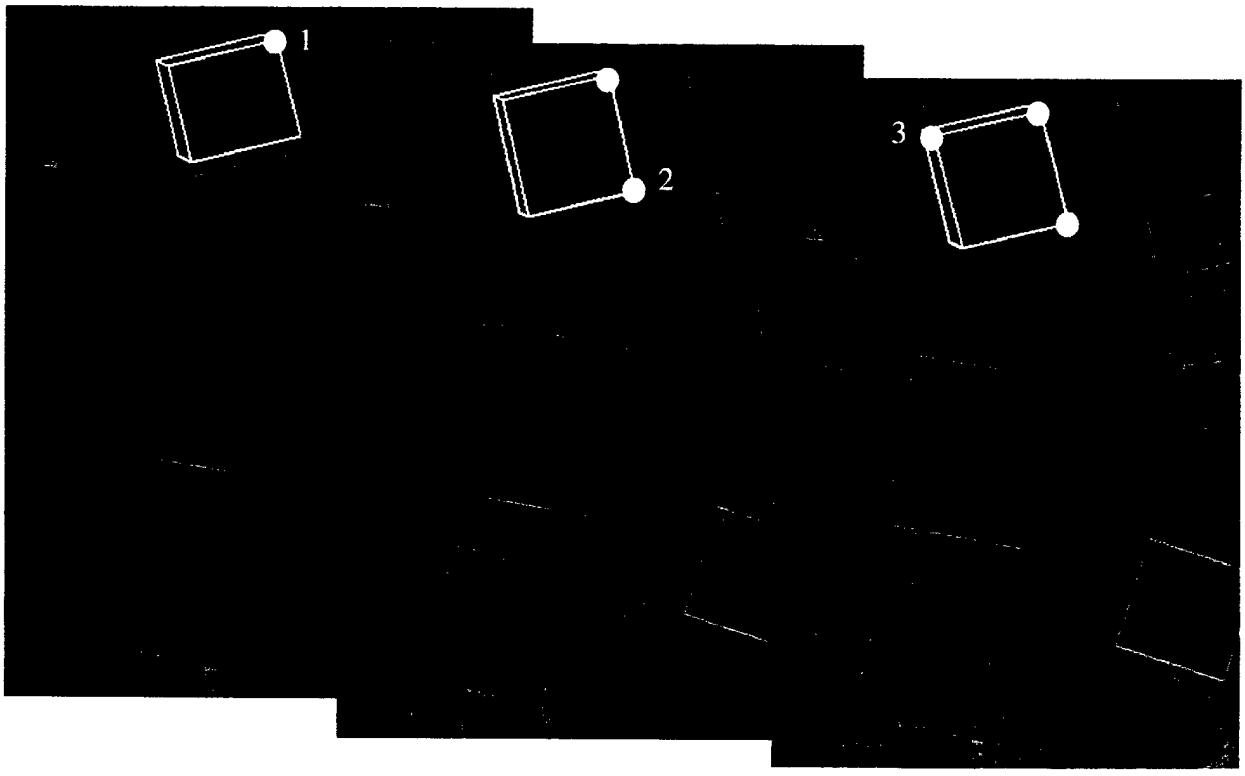


Figure 4.5 Adding this building required three clicks

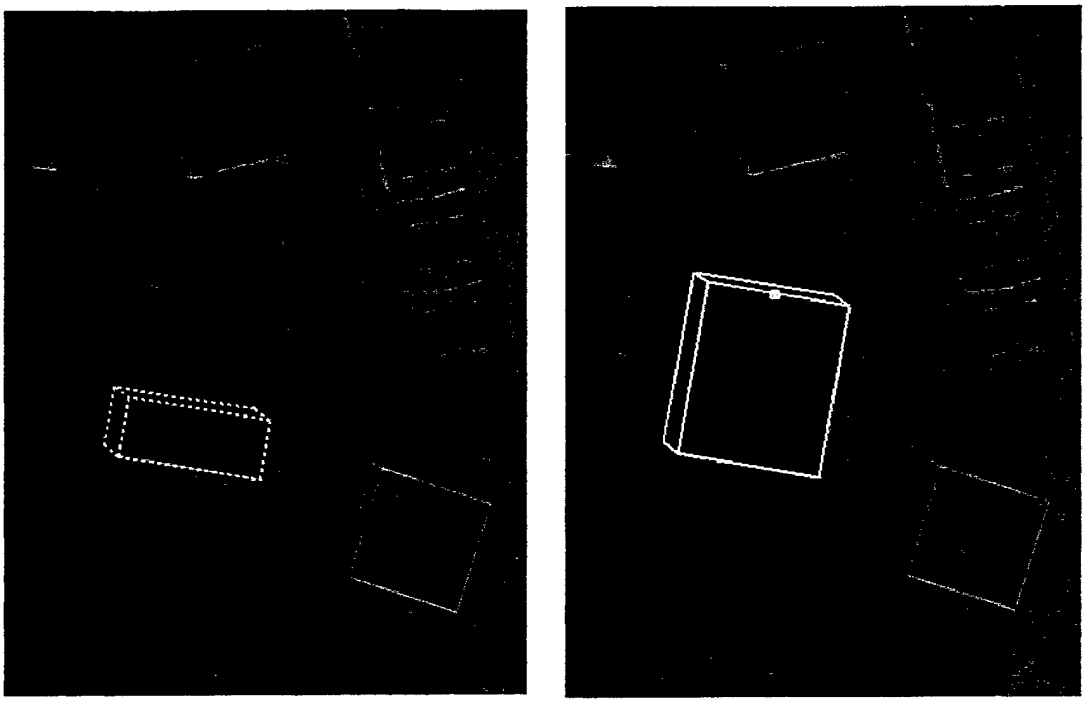


Figure 4.6 Adjusting the location of one side requires one click anywhere along the correct boundary of the roof.

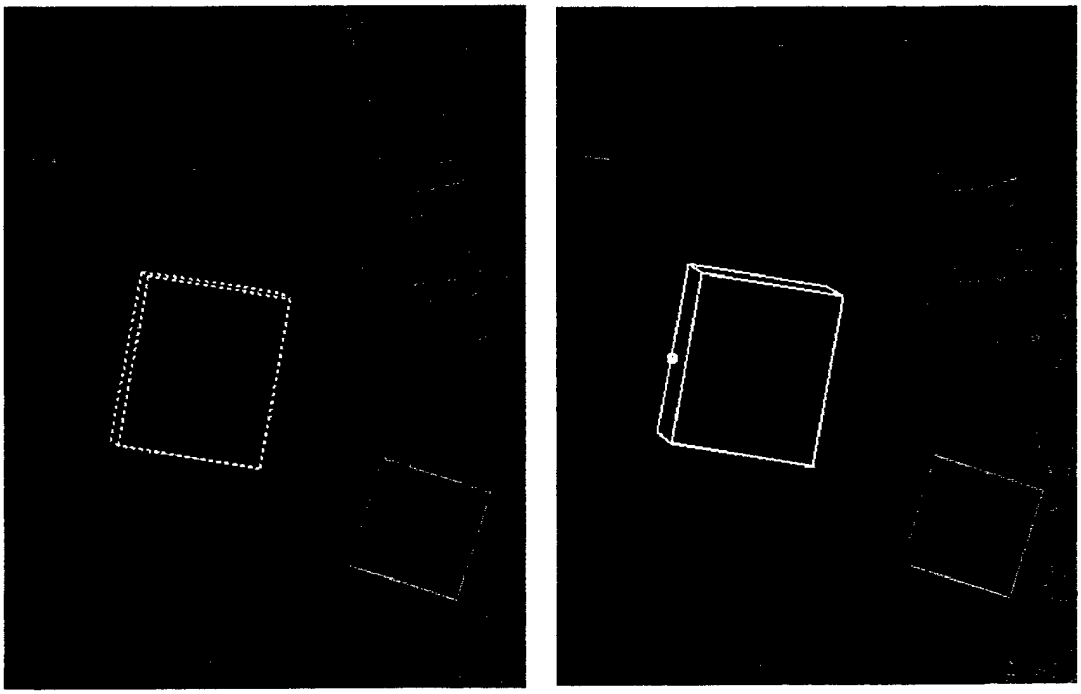


Figure 4.7 Adjusting the height of a building requires selection of a point anywhere along the base of the building.

5 Results, Conclusion and Future Work

The system described in Section 3 has been applied to several windows of the Fort Hood site and to the part of McKenna MOUT site at Fort Benning that contains flat roofs. Space does not permit including them all in this report. Instead, some selected examples are shown. The examples shown are of relative more complex buildings. The results shown are indicative of the overall performance.

A formal quantitative evaluation of the result is yet to be completed. The evaluation metric and methodology is in the process of being defined by the Integrated Feasibility Demonstration (IFD) contractor and other participants in this program. A batch-mode version of this system has been ported to the IFD contractor for integration and evaluation into a larger system.

A qualitative discussion of results is provided below. Use of IFSAR is illustrated where available. All examples were run with the same parameter settings and not all the examples were used in the training of the verification steps.

Figure 5.1 shows results on an area of Fort Hood with complex shaped buildings. Two of the buildings in the lower part are not detected. Most others are detected quite accurately though dimensional errors exist in the smaller parts. The confidence value are shown color coded as before (red is high, orange is medium and yellow is low). Most false alarms are of low confidence, however, removing them by raising the acceptance threshold will also eliminate some of the actual buildings. Use of IFSAR cues, in the verification stage, can also help eliminate false alarms; these results are shown in Figure 5.2. All false alarms are removed, however, one correct building has also been removed because it is not strongly represented in the IFSAR image. Note that the missing buildings are also not salient in the IFSAR image.

Figure 5.3 shows results on part of the McKenna MOUT site at Fort Benning. The four, large flat roof buildings are detected accurately and with high confidence. Some flat roof structures are also detected as part of gable buildings; these should disappear when gable processing is completed. Their is only one structure, shown in orange, that does not correspond to a building. This is easily eliminated by using the IFSAR cue for verification.

Figure 5.4 shows results on the "headquarters" part of the Fort Hood site by using a different pairs of images then in results shown earlier in Section 3; one of the images in the pair is a near-nadir image. The results are better than before and the large bright building in the lower left part is now detected, though its dimensions are somewhat incorrect. Good IFSAR data was not available for this part of the scene.

Figure 5.5 shows yet another part of Fort Hood with several complex buildings. The results are again similar to those for other areas.

Future plans for this project include further development of slanted roof detection and delineation, more integrated use of IFSAR data, improvements in the selection and verification processes by using ore complex decision methods, thorough quantitative evaluation on available datasets and extension to more complex class of buildings as data becomes available.

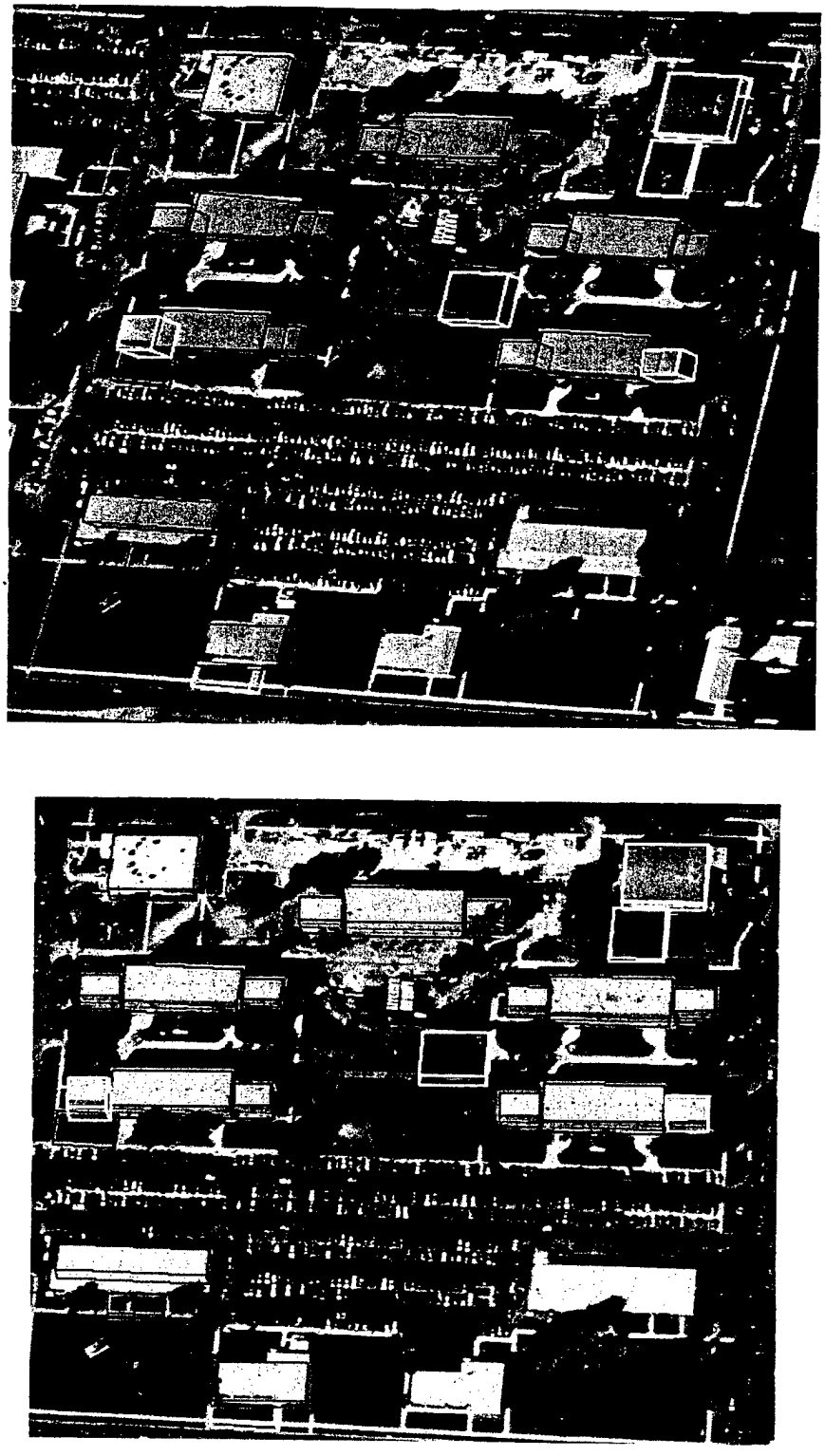


Figure 5.1 Result of processing a portion of two different views of Fort Hood without use of IFSAR cueing

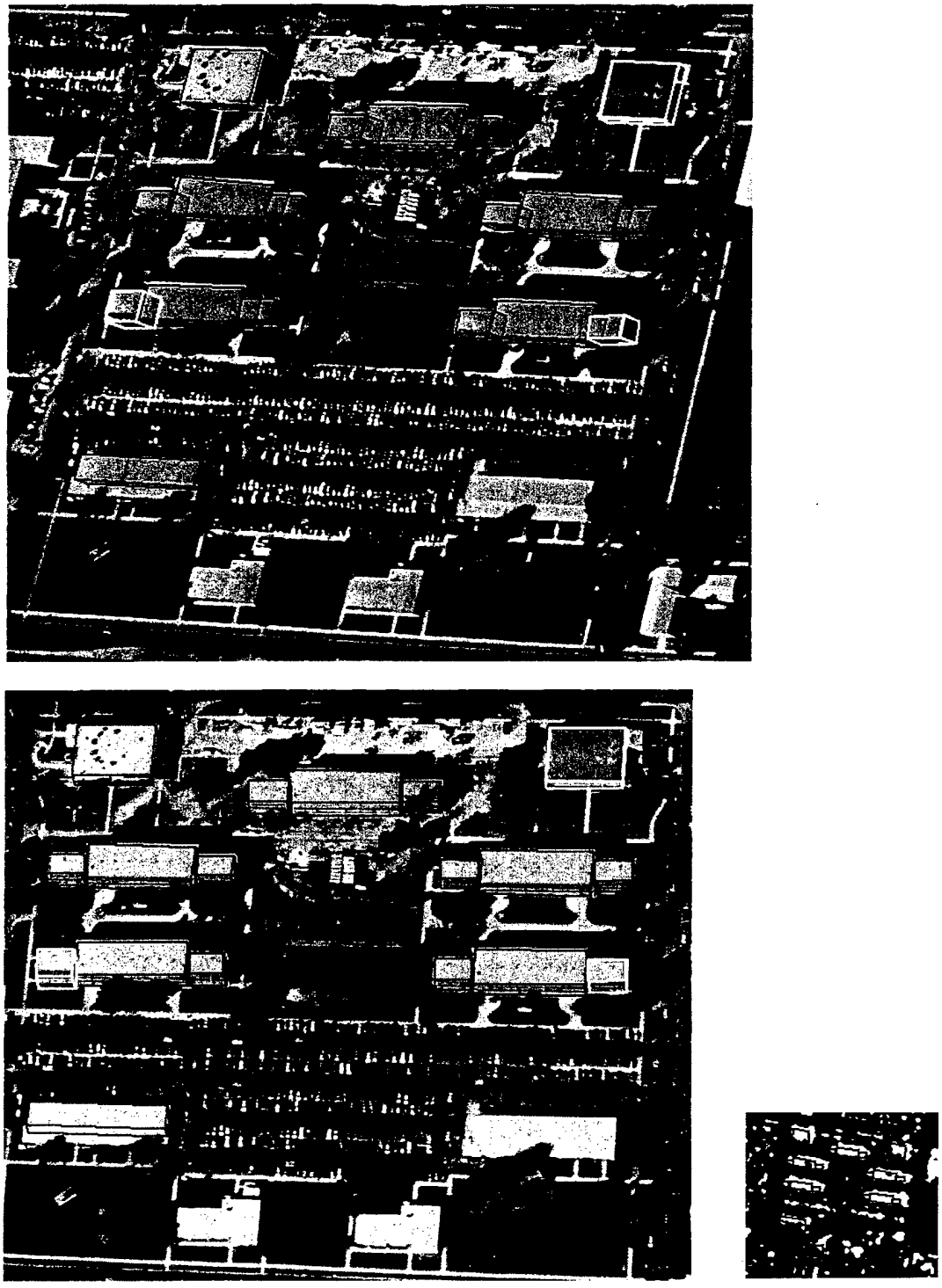


Figure 5.2 Result for the same portions of Fort Hood using IFSAR cueing

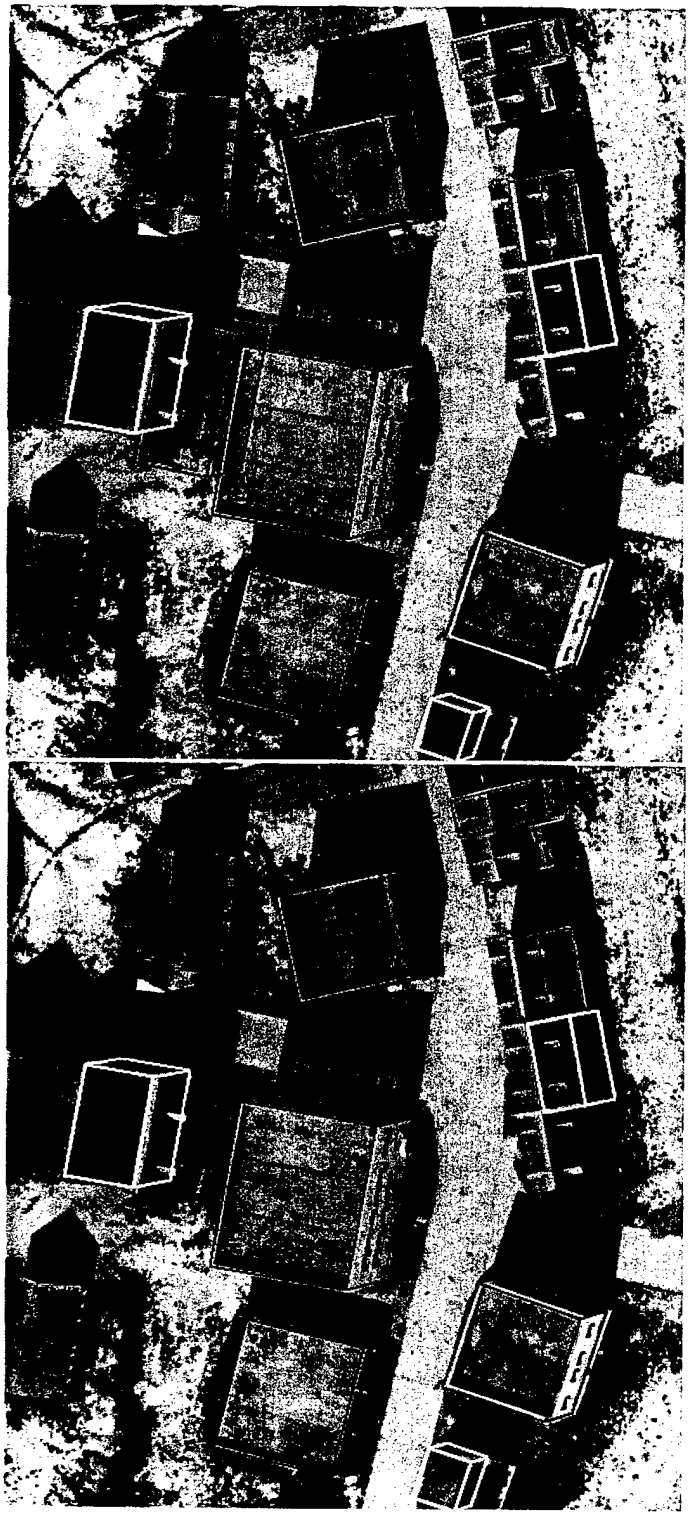


Figure 5.3 Results for a portion of the McKenna MOUT without (top) and with IFSAR cueing

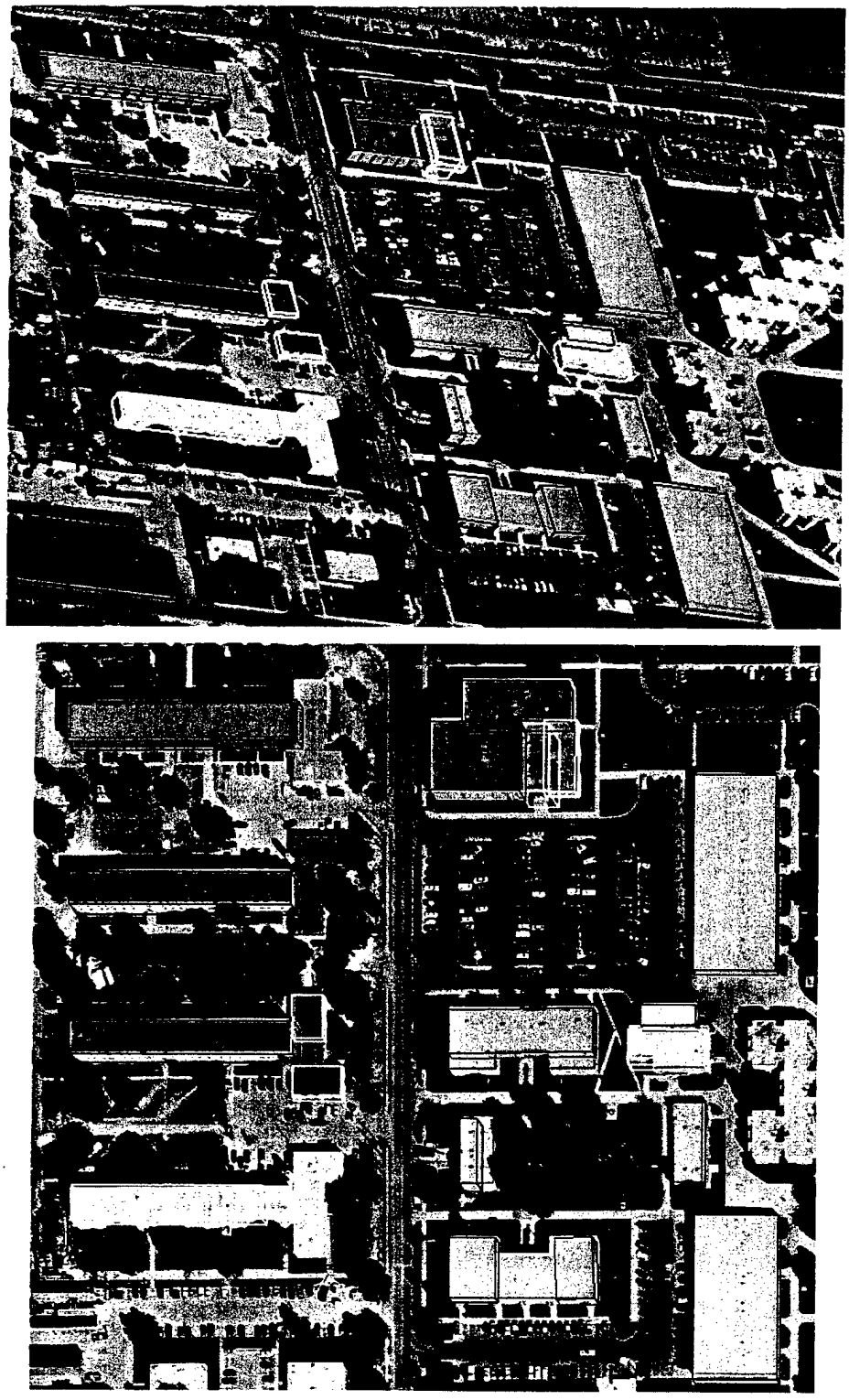


Figure 5.4 Results for Headquarters area of Fort Hood without IFSAR cueing

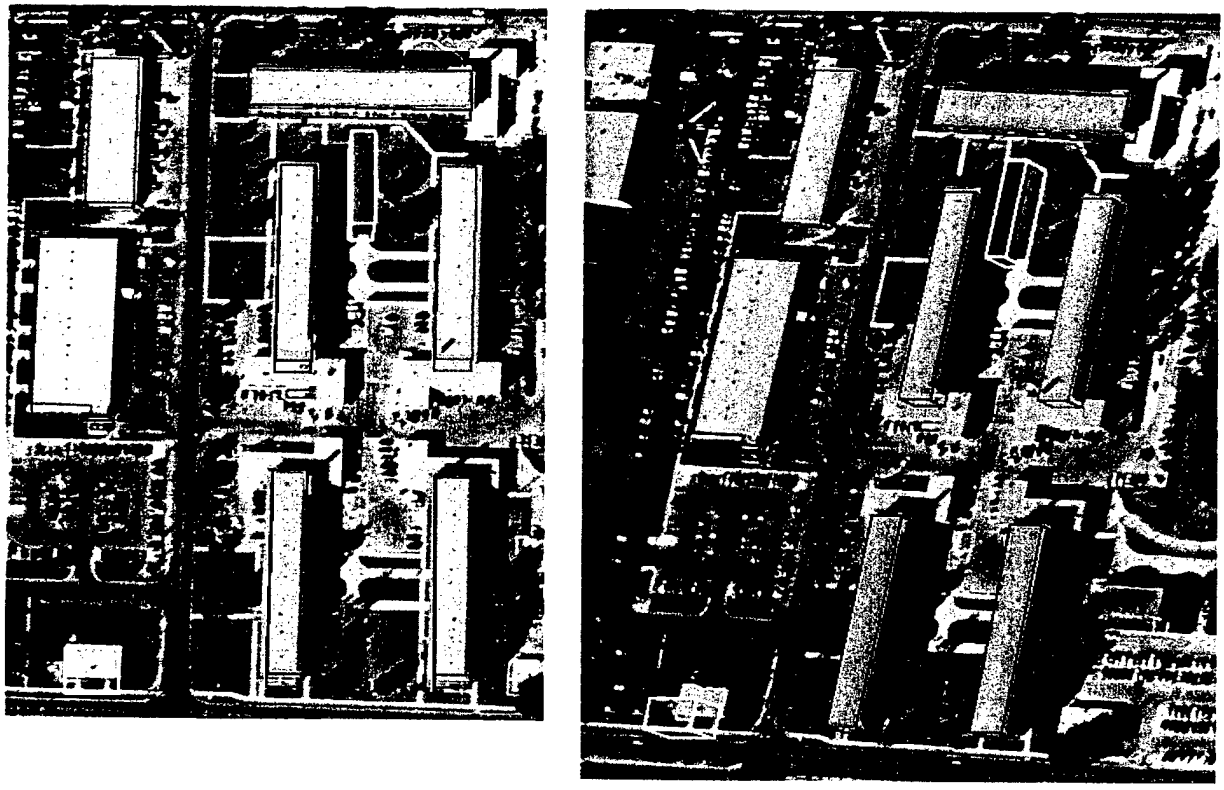


Figure 5.5 Result for another area of Fort Hood without IFSAR cueing

6 References

[Chen *et al.*, 1987] J. Chen, A. Huertas and G. Medioni, "Fast Convolution with Laplacian-of-Gaussian Masks", IEEE Transactions on Pattern Analysis and Machine Intelligence", Vol. PAMI-9, No.4., July 1987, pp 585-590.

[Chung & Nevatia, 1992] C.-K. R. Chung and R. Nevatia, "Recovering LSHGCs and SHGCs from Stereo," In *Proceedings of the DARPA Image Understanding Workshop*, San Diego, CA, January 1992, pp. 401–407.

[Gruen & Baltsavias, 1997] A. Gruen and M. Baltsavias, Editors, "Automatic Extraction of Man-Made Objects from Aerial and Space Images", Birkhauser Verlag, 1997.

[Heuel & Nevatia, 1995] S. Heuel, R. Nevatia, "Including Interaction in an Automated Modeling System", Symposium on Computer Vision, Miami Beach, FL, November 1995(383-388).

[Heuel & Nevatia, 1996] S. Heuel and R. Nevatia, "Including Interaction in an Automated Modeling System," in *Proceedings of Image Understanding Workshop*, Palm Springs, CA, February 1996, pp. 429-434.

[Huertas & Medioni, 1986] A. Huertas and G. Medioni, "Detection of Intensity Changes with Subpixel Accuracy using Laplacian-of-Gaussian Masks", IEEE Transactions on Pattern Analysis and Machine Intelligence", Vol. PAMI-8, No.5., September 1986, pp 561--664.

[Huertas, *et al.*, 1990] A. Huertas, W. Cole, and R. Nevatia, "Detecting Runways in Complex Airport Scenes," *Computer Vision, Graphics, and Image Processing*, 51(2):107–145, August 1990.

[Huertas, *et al.*, 1995] A. Huertas, M. Bejanin and R. Nevatia. "Model Registration and Validation", in *Proceedings of the Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Ascona, Switzerland, Birkhaiser, March 1995, pp 33-44.

[Huertas & Nevatia, 1997] A. Huertas and R. Nevatia. "USC Radius Related Research: An Overview", in Proceedings of Image Understanding Workshop, New Orleans, LA, May 1997, pp. 437-447.

[Huertas & Nevatia, 1998] A. Huertas and R. Nevatia, "Detecting Changes in Aerial Views of Man-Made Structures" in *Proceedings of International Conference on Computer Vision,* pp.73-79, Bombay, India, January, 1998.

[Ito & Ishii, 1986] M. Ito and A. Ishii, "Three-view stereo analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, 8:524-532, 1986.

[John & Langley, 1995] George H. John and Pat Langley, "Estimating Continuous Distributions in Bayesian Classifiers", Proc. 11th Conf. on Uncertainty in Artificial Intelligence, San Mateo, CA 1995

[Lin *et al.*, 1995] C. Lin, A. Huertas, and R. Nevatia, "Detection of Buildings from Monocular Images", in *Proceedings of the Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Ascona, Switzerland, pp 125-134, March 1995.

[Lin & Nevatia, 1995] C. Lin, and R. Nevatia, "3D Descriptions of Buildings from an Oblique View Aerial Image", IEEE International Symposium of Computer Vision, 377-382, 1995.

[Lin et al., 1997] C. Lin, A. Huertas, and R. Nevatia, "A System for Building Detection from Aerial Images" in *Proceedings of the Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images II,* Ascona, Switzerland, pp 77-86, March 1997.

[Medioni & Nevatia, 1984] G. Medioni and R. Nevatia, "Matching Images Using Linear Features," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):675–685, November 1984.

[Medioni & Nevatia, 1985] G. Medioni and R. Nevatia, "Segment-Based Stereo Matching," *Computer Graphics and Image Processing*, 31(1):2–18, July 1985.

[Medioni et. al., 1991] G. Medioni, A. Huertas and M. Wilson, "Automatic Registration of Color Separation Films", Machine Vision and Applications Journal, Springer-Verlag, NY, Vol. 4, 1991, pp. 33-51.

[Mohan & Nevatia, 1989a] R. Mohan and R. Nevatia, "Perceptual Organization for Segmentation and Description," in *Proceedings of the DARPA Image Understanding Workshop*, Palo Alto, California, May 1989. Morgan Kaufmann Publishers, Inc.

[Mohan & Nevatia, 1989b] R. Mohan and R. Nevatia, "Segmentation and Description Based on Perceptual Organization," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 333–341, San Diego, California, June 1989.

[Mohan & Nevatia, 1989c] R. Mohan and R. Nevatia, "Using Perceptual Organization to Extract 3-D Structures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(11):1121–1139, November 1989.

[Noronha & Nevatia, 1997] S. Noronha and R. Nevatia, "Detection and Description of Buildings from Multiple Aerial Images," in *Proceedings of the DARPA Image Understanding Workshop*, New Orleans, LA, May 1997

[Noronha & Nevatia, 1997b] S. Noronha and R. Nevatia, "Detection and Description of Buildings from Multiple Aerial Images", IEEE Computer Vision and Pattern Recognition, pp. 588-594, San Juan, PR, June, 1997.

[Noronha & Nevatia, 1998] S. Noronha and R. Nevatia, "Recent Advances in Detection and Description of Buildings from Multiple Aerial Images", Asian Conference on Computer Vision, Hong Kong, January, 1998.